Seed sourcing in the genomics era

Colette Blyth

This thesis is presented in fulfilment of the requirements for the degree of

Doctor of Philosophy

The University of Adelaide
School of Biological Sciences
July 2021

Blank page

Table of Contents

Publications included in this thesis

1. Blyth, C., Christmas, M.J., Bickerton, D.C., Faast, R., Packer, J.G., Lowe, A.J. and Breed, M.F., 2020. Increased genetic diversity via gene flow provides hope for Acacia whibleyana, an endangered wattle facing extinction. *Diversity*, *12*(8), p.299. – **Chapter 2.**

2. Blyth, C., Breed, M.F., Hodgson, R. J., Gellie, N, J. C., Lowe, A.J, McDonald, J., Christmas, M.J., In Prep. Population genomics for delineating rangewide seed zones for restoration of a keystone arid-zone shrub. *In preparation for submission to Restoration Ecology*. – **Chapter 3.**

3. Blyth, C., Christmas, M.J., Bickerton, D.C., Breed, M.F., Foster, N.R., Guerin, G.R., Mason, A.R. and Lowe, A.J., 2021. Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy. *Life*, *11*(6), p.553. – **Chapter 4.**

Abstract

How do we restore and conserve plant species in a drastically changing world? Sourcing enough high-quality seed is one of the major challenges facing practitioners, and meeting this demand is complicated by habitat fragmentation decreasing the genetic diversity of populations and climate change shifting the adaptive landscape. The traditional method of local seed sourcing does not account for these issues and, therefore, there has been much discussion over suitable alternatives. Fundamentally, a solid understanding of the levels of differing levels of genetic diversity within and divergence between the populations of a target species is essential. In this thesis I use next generation sequencing (NGS) as a tool to increase our understanding of genetic structure and environmental adaptation to make informed seed sourcing decisions aimed at establishing more resilient plant populations to climate change. Rangewide NGS datasets for three native Australian species provide case studies for the integration of genomics into restoration and conservation decision making. I generate and analyse NGS, phenotypic, and environmental datasets to inform the unique set of management approaches required for each species, with the overall aim of achieving successful conservation and restoration of Australian plant species.

The three core studies of this thesis consist of:

1) The conservation management of the endangered Whibley wattle (*Acacia whibleyana*).

2) Revegetation strategies for the widespread and abundant pearl bluebush (*Maireana sedifolia*), which is heavily used in post-mining rehabilitation.

3) Species boundaries and conservation management of the wide-ranging yet vulnerable daisy bush (*Olearia pannosa*).

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree. I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Research Council Funded Scholarship.

Colette Blyth

Acknowledgements

Firstly, to home.

To my parents — Mum, Paul, Dad, and Kate. I hope that this makes you proud. I haven't once doubted all of your love and support, despite the distance and the time apart. A little extra thanks to my dad, Stuart, for all your help with drafting and always holding me to high standards. To my wee brother Sam — if I can do a PhD, you've no excuse. Thanks for your help with the tedious things and the pretentious music recommendations. To my baby sister Dani, who makes me proud every day. I've watched you single-handedly raise a child, struggle with your health, and establish a successful career. You put it all into perspective and you're one of my heroes.

To Hannah Blyth, a true constant in my life since Betty Hill.

To when the borders reopen.

I also want to give a special acknowledgement to Matt Christmas. I have absolutely no doubt that the submission of this thesis is a result of your consistent and patient hands-on supervision alongside your warmth and support in all the unexpected, non-research related difficulties of my PhD. I think we spent less than 24 hours in the same room during my studies, yet I felt like I shared an office with you for the duration of them. COVID-19 scuppered our plans last year, but I will be buying you a beer when the world goes back to normal!

Next, thanks go to the rest of my supervisory panel. To Jeremy Austin for coming in late to steer me through and to Andy Lowe for providing my PhD with flair and support whenever I asked for it. Thanks should also go to Martin Breed, this whole thing was your idea, after all! I would also like to thank my postgraduate coordinator Dr Steve Delean for providing me with support behind the scenes.

I would also like to thank all the co-authors, collaborators, and partners I have worked alongside at the University of Adelaide, Terrestrial Ecosystem Research Network, the Department of Environment and Water, Trees for Life, Iluka Resources Ltd., The State Herbarium of South Australia, and the academics and researchers who have provided me advice and input along the way. I've been met with nothing but kindness and enthusiasm, and I've learned from you all. Also, to all my volunteers who helped out with fieldwork, lab work, and various components of the analysis and writing process: Elleca Polson, Sanjaya Rai, Dona Kireta, Romy Dresken, Harrison Sydnor, Jacob Mills. Particularly, to Kiri Marker who joined me as a volunteer and three tumultuous years on is my best friend and partner. Thanks for cooking tonight so I can write this.

As an international student in Adelaide, it was hard to find my place in a less transient social environment than I was used to, with homesickness heightened by the impacts of COVID-19. So lastly, but perhaps most importantly, I would like to thank the community that embraced me. To James Tresize — my oldest friend here, whose family now treats me as one of their own. To Jacob Mills who has shared a special part of this country with me, and many a whisky too. To Nick Gellie, for providing support, wisdom, and excellent conversation during some very difficult times. To Adam Toomes, we ran the Biology Society together for a year, but we also shared much more than that. To Stefan Caddy-Retalic, Emrys Leitch and Benny. Brown Hill Creek was magnificent for so many reasons. Thank you for sharing it.

To Nicole Foster – the friend I met at precisely the right time. Thank you for the past year. It's been nothing but an adventure. We did it!

To all of the music, books, art, poetry and things that kept me going when it all felt too much or not enough.

To nature.

The Darkling Thrush

I leant upon a coppice gate
When Frost was spectre-grey,
And Winter's dregs made desolate
The weakening eye of day.
The tangled bine-stems scored the sky
Like strings of broken lyres,
And all mankind that haunted nigh
Had sought their household fires.

The land's sharp features seemed to be
The Century's corpse outleant,
His crypt the cloudy canopy,
The wind his death-lament.
The ancient pulse of germ and birth
Was shrunken hard and dry,
And every spirit upon earth
Seemed fervourless as I.

At once a voice arose among
The bleak twigs overhead
In a full-hearted evensong
Of joy illimited;
An aged thrush, frail, gaunt and small,
In blast-beruffled plume,
Had chosen thus to fling his soul
Upon the growing gloom.

So little cause for carolings
Of such ecstatic sound
Was written on terrestrial things
Afar or nigh around,
That I could think there trembled through
His happy good-night air
Some blessed Hope, whereof he knew
And I was unaware.

Thomas Hardy, 1900

Chapter 1. Introduction

Aspects of this chapter were included in the opinion article in Nature Reviews Genetics — 'The potential of genomics for restoring ecosystems and biodiversity' (Breed et al., 2019).

Land degradation

The human 'footprint' on the planet now covers approximately 83% of the earth's land surface (Hooke et al., 2012, Sanderson et al., 2002). A major consequence of this occupation is land degradation, which has been driven largely by broadscale clearing of native vegetation. Land degradation is defined as the "long term decline in ecosystem function and productivity" (Bai et al., 2008), it is estimated that approximately one third of all arable land is now degraded (Gibbs and Salmon, 2015, Nkonya et al., 2016). Direct consequences of land degradation are habitat loss and fragmentation, which are causing dramatic declines in biodiversity (Nkonya et al., 2016, WWF, 2016). A recent survey of global biodiversity indicates a 58% global decline in wildlife populations since 1970, with land degradation and habitat loss being the most common cause (WWF, 2016).

While there are ethical concerns regarding the preservation of biodiversity, loss of biodiversity also has direct implications for human populations, including the loss of ecosystem services. Defined as "the benefits people derive, directly or indirectly, from ecosystem functions" (Costanza et al., 1997), ecosystems provide many essential requirements of human societies including the pollination of food crops, carbon sequestration and oxygen production of forests, clean water provision, and products ranging from timber to medicines (Costanza et al., 1997, Fargione et al., 2008, Fisher and Turner, 2008, Kennedy et al., 2013). The continued supply of these services is dependent on the reduction in the rate at which land degradation is occurring, the conservation of existing habitat, and on the restoration of degraded land (Folke et al., 2004).

Conservation of rare and threatened plant species

In order to persist, plant populations must overcome a multitude of threats. Pests, invasive species, disease, a rapidly changing climate, land clearance and habitat fragmentation, and genetic consequences of small population size have been identified as the major threats faced by plant species by Oostermeijer (2003). In a relatively short period of time, the native vegetation of Australia has either been extensively cleared or adopted for stock grazing (e.g., rangelands) (Broadhurst and Young, 2007, Harrington et al., 1984). In addition to the threats outlined by Oostermeijer (2003), Australian plant populations experience additional threats from overgrazing by stock, feral animals (e.g., goats) and fire (Broadhurst and Young, 2007, Raven and Yeates, 2007). In order to contend with these threats, practitioners must decide on the correct management interventions on a case-by-case basis, weighing up the costs and benefits of various methods. In some instances, the method employed will be direct, such as the removal of invasive species (see Hoffman and Broadhurst, 2016), or fire management (e.g., Taylor et al., 2013). There is also increasing focus on conserving genetic diversity in populations to increase their evolutionary potential and so improve their resilience to future change (Sgrò et al., 2011). The key points of this are discussed below, but put simply, the core tenets of conservation genetics state that if populations become small and isolated, they are threatened by genetic erosion caused by genetic drift (the natural fluctuations in allele frequencies over time through random sampling, which often leads to loss of genetic diversity in small populations) (Frankham et al., 2009, Reed and Frankham, 2003). Genetic drift can drive genetic divergence between populations and cause the loss of adaptively important alleles, or the fixation of non-advantageous alleles. This affects both the short-term and long-term genetic health of a population, through both the reduction in individual fitness and a reduction in the overall adaptability of a population to future change (Frankham et al., 2009). Furthermore, inbreeding depression can occur in small populations as a result of a higher likelihood of reproduction between related individuals and can

accelerate genetic drift (Frankham et al., 2009). To address this, an increasingly common solution is to facilitate gene flow between populations to provide genetic rescue. Genetic rescue theoretically bolsters the genetic diversity within a population and increase its likely adaptability to change. In plants, this is carried out by translocations. *The Guidelines for the Translocation of Threatened Plants in Australia* (Commander et.al., 2018) define translocations as "The deliberate transfer of plants or regenerative plant material from an ex-situ collection or natural population to a new location, usually in the wild. It includes reintroduction, introduction, reinforcement, assisted migration and assisted colonization." Put simply, a translocation is the act of moving individuals within a species from one location to another for conservation or restoration purposes. The primary goals can be the establishment or re-establishment of a population, to increase population size through augmentation, or to bolster genetic diversity introducing new genetic material (Commander et.al., 2018). There are a variety of translocation methods, ranging from transportation of plants, seeds, and propagules between sites to the ex-situ propagation of plants for future planting. For a translocation programme to be considered a success, a population must be deemed self-sufficient in the long-term (Commander et.al., 2018). The key factors attributed to the success or failure of a translocation programme are: planting enough individuals to ensure adequate survival rates, controlling threats (e.g., weeds), site suitability, and long-term maintenance (Commander *et al.* 2018). Examples of successful translocation efforts can be found in the endangered *Acacia whibleyana* (a species studied in this thesis). Due to population decline and issues with recruitment, practitioners collected seed from remnant stands and grew them in nurseries to plant back out as seedlings to augment the population that they were collected from. This was considered a success as although just under 50% of the seedlings established, this still resulted in a 25% increase in population size of the target population (Jusaitis and Sorenson, 2007). Using the same propagation and method, a new, translocated population was later successfully established in a roadside reserve through two

planting events. As of the 2019, there was estimated to be 42 individuals established in this population (Faast et. al, 2019).

Ecological restoration

Ecological restoration is a rapidly growing field (Shackelford et al., 2013). Initiated as a direct response to land degradation, it is defined as the "process of assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed" with a primary goal of aiding the recovery of lost biodiversity and ecosystem function in affected areas (SER, 2004). Due to growing concerns over the impact of land degradation, there has been large scale investment in restoration projects (Breed et al., 2013, Feng et al., 2013). Initiatives range from small-scale, community-based projects (e.g., Caledonian woodland in Scotland) (BBC, 2016) to 'super-sized' programs (e.g., Chinese government's $US40 billion 'Grain to Green' program) (Feng et al., 2013). The Bonn Challenge is perhaps the most significant development in restoration, and is a global effort backed by the 2014 United Nations Climate Summit, which, in partnership with governments, companies, and communities, aims to restore 350 million ha of cropland and forest by 2030 (Minnemeyer et al., 2011, Verdone and Seidl, 2017). So far, 156 million ha of land has been committed for restoration as part of the Bonn Challenge and it has been estimated that, if fully achieved, the Bonn Challenge will generate a net economic benefit of $US 9 trillion through ecosystem services (Verdone and Seidl, 2017).

Often, a major step in the process of ecological restoration is revegetation (Clewell and Rieger, 1997, Crouzeilles et al., 2016). Historically, however, success rates have been low (Crouzeilles et al., 2016). A recent global meta-analysis by Crouzeilles et al. (2016) reported high variation in woodland restoration success. Post monitoring of restoration projects recorded a median biodiversity enhancement range of 15-84% and a median vegetation structure enhancement range of 36-77%. These success rates reflect the results of Godefroid et al. (2011), who estimated average survival rates of

plant reintroductions to be as low as 50%. Poor seed quality is a fundamental cause of the poor success rate found in revegetation projects (Breed et al., 2013, Broadhurst et al., 2008, Godefroid et al., 2011).

Seed sourcing for revegetation

Traditionally, revegetation projects have relied on harvesting seed from local populations, assuming local seed to be better adapted to local conditions compared to seed sourced elsewhere (Breed et al., 2013, Broadhurst et al., 2008). However, local seed-sourcing can have several limitations. Firstly, populations local to restoration sites are often fragmented and small. Such populations typically suffer lowered genetic diversity, associated with increased genetic drift (Honnay and Jacquemyn, 2006, Vranckx et al., 2011, Young et al., 1996), and elevated inbreeding rates (Breed et al., 2012, Breed et al., 2013, Eckert et al., 2010, Lowe et al., 2015). As such, seed from these populations is likely to be of poor quality. Furthermore, if the source populations are small and/or fragmented, they are often vulnerable to overharvesting, which raises concerns over the sustainability of this method (Broadhurst et al., 2015, Broadhurst et al., 2008). Secondly, we are now in an age of accelerated climate change (Cook et al., 2013, Pachauri et al., 2014, Parmesan, 2006), with implications for plant life (Scheffers et al., 2016). When confronted with changing conditions, plants can respond in one of three ways: adapt to the new environment, migrate to an area with suitable conditions, or perish (Aitken et al., 2008, Christmas et al., 2015). Conditions are now changing so rapidly that plants may not have the capacity to respond through migration or adaptation (Jump and Peñuelas, 2005). Consequently, practitioners using local seed in areas of radically changing environment may simply be introducing maladapted progeny with low survival rates. In light of these concerns, there has been much debate over suitable provenancing for seed sourcing and a shift away from the 'local is best' paradigm.

For revegetation projects to succeed, access to large quantities of high-quality seed is essential (Breed et al., 2013, Broadhurst et al., 2008, Godefroid et al., 2011). A key step in this process is choosing where to source this seed from (Breed et al., 2013, Broadhurst et al., 2008, Mijnsbrugge et al., 2010, Rice and Emery, 2003). Typically, this has involved collecting seeds from plant populations thought to be well adapted to the target restoration site (Breed, 2013, Broadhurst et al., 2008). Historically, in Australia, collection has tended to occur within a proximity of <15 km to the restoration area based upon the home-site advantage hypothesis (Kawecki and Ebert, 2004, Montalvo and Ellstrand, 2000). This hypothesis assumes that plants are locally adapted on a small scale and, therefore, collecting local seed minimises any risks associated with introducing novel genotypes into an area (Broadhurst et al., 2008, Mortlock, 2000a, Mortlock, 2000b). These risks include lowered fitness due to maladaptation; outbreeding depression due to hybridisation of non-local and local genotypes; 'invasive genotypes' causing the introduced population to become overly dominant within an ecosystem, resulting in negative effects on the ecosystem as whole (Mijnsbrugge et al., 2010). These risks have been considered such important factors in seed sourcing guidelines that species have been omitted from restoration projects if seed could not be sourced locally (Broadhurst et al., 2008).

While concerns have been raised over the potential for genetic contamination when restoring plant populations with non-local populations (Hufford and Mazer, 2003, Krauss et al., 2013, McKay et al., 2005), current evidence has clearly refuted the assumption that all plant populations are locally adapted. Both a meta-analysis by Leimu and Fischer (2008) and a study by Hereford (2009) found that small-scale adaptation is not universal. Indeed, these meta-analyses ruled out geographic distance as a predictor of local adaptation, with the strongest predictor of local adaptation being population size, with larger populations demonstrating more local adaptation (Leimu and Fischer, 2008).

Founder effects can play large roles in the success or failure of restoration efforts, especially where a species is being reintroduced. Founder effects occur when a population is established by too few individuals, which together harbour only a small portion of the genetic diversity found in the wider population, resulting in a genetic bottleneck (Frankham, 2005). If seed is sourced only from local populations, which may already be genetically compromised by fragmentation and habitat loss, then genetic diversity will be low, potentially jeopardising the success of the restoration efforts by seed sourcing causing a genetic bottleneck (Hufford and Mazer, 2003).

The 'local is best' approach also rests on the assumption that environmental factors at the restoration site have remained constant through time, thus allowing local populations to be suitably adapted (Wilkinson, 2001). A changing climate creates a shifting adaptive landscape. Further, a negative synergy between habitat fragmentation and climate change is a concerning prospect — in the face of rapid climate change, fragmented populations may find themselves stranded in areas where local conditions have shifted away from the population's optimum and gene flow of adaptive alleles from other populations is heavily restricted due to their isolation (Breed et al., 2013, Havens et al., 2015, Wilkinson, 2001). Widespread evidence of local maladaptation has already been noted. Maschinski and Haskins (2012) found that progeny of the flowering plant, beach clustervine (*Jacquemontia reclinate*), performed best in reintroductions when seed was produced from mixed sources rather than local only. Not only did the mixed source progeny outperform their local counterparts, they proved to be more resilient to extreme climate events. Similarly, the work of Gellie et al. (2016) found local maladaptation in a transplant trial of the tree species *Eucalyptus leucoxylon*. The study tested for local adaptation and sampled from three provenances across an aridity gradient. They found the local provenance to perform poorest in plant height, survival rates, and pathogen resistance, and recommended using more diverse seed mixes in future restoration projects. As evidence directs seed sourcing

away from local collections, the focus has shifted to establishing resilient populations in the face of habitat fragmentation and climate change (Aitken and Whitlock, 2013, Breed et al., 2013, Broadhurst et al., 2008, Gellie et al., 2016, Prober et al., 2015, Ramalho et al., 2017).

The IUCN identified genetic diversity as one of the three levels of biodiversity to be conserved (Reed and Frankham, 2003). If low levels of genetic diversity are present, then a population will have a reduced capacity to adapt. Population fitness has been directly correlated with genetic diversity. For example, in a meta-analysis that included vertebrates, invertebrates, and plants, Reed and Frankham (2003) found genetic diversity to explain 19% of fitness variation. In plants, genetic diversity and therefore fitness, is primarily associated with population size, ecological genetics, and mating system (Ellegren and Galtier, 2016). Genetic diversity positively correlates with population size, as a reduction in population size limits mate availability and increases inbreeding and the risk of genetic drift (Leimu et al., 2006). Plant mating systems range from complete outcrossing to selfing, with mixed mating commonly observed (Imai et al., 2016). As new genetic variation (i.e., that which arises though de novo mutations) is most commonly distributed between plant populations though outcrossing, genetic diversity is generally higher in species which adopt this mating system. (Hamrick and Godt, 1996).

An early departure from recommendations of strict local seed sourcing can be found in Lesica and Allendorf (1999), who concluded that local was only best if sites were not heavily degraded. Instead, they advocated genetic mixing for such sites to create more resilient populations, a recommendation later reiterated in a meta-analysis by Godefroid et al. (2011). Wilkinson (2001) raised concerns over maladaptation of the local provenance due to habitat fragmentation and changing environmental conditions. Seminal to the shift away from local adaptation was the work by Broadhurst et al. (2008) with their 'composite provenancing' approach. Composite provenancing is a seed collecting method that recommends seed is

collected mainly from the local provenance but is then mixed with seed sourced from populations found along a scale of increasing distance from the restoration site. The objective being to replicate natural long-distance gene flow, albeit over a much shorter timeframe, theoretically enhancing the adaptive capacity of a population. Several additional alternatives to local seed sourcing have since been discussed in the literature, and it is possible to categorise the majority of techniques into one of two broad groups – those that focus on mixing seed to increase overall adaptive potential of a population and those that attempt to provide the genotypes predicted to match projected change in environmental condition (see Figure 1. for summary of seed sourcing methods; Breed et al., 2013, Broadhurst et al., 2008, Prober et al., 2015, Ramalho et al., 2017, Sgrò et al., 2011, Williams et al., 2014).

Many approaches found in restoration and conservation programmes share the common theme of how best to facilitate and/or establish genetically diverse and resilient populations capable of adapting to change. The application of these methods is underpinned by an understanding of the underlying genetic diversity and divergence of the restoration species (Breed et al., 2019, Williams et al., 2014).

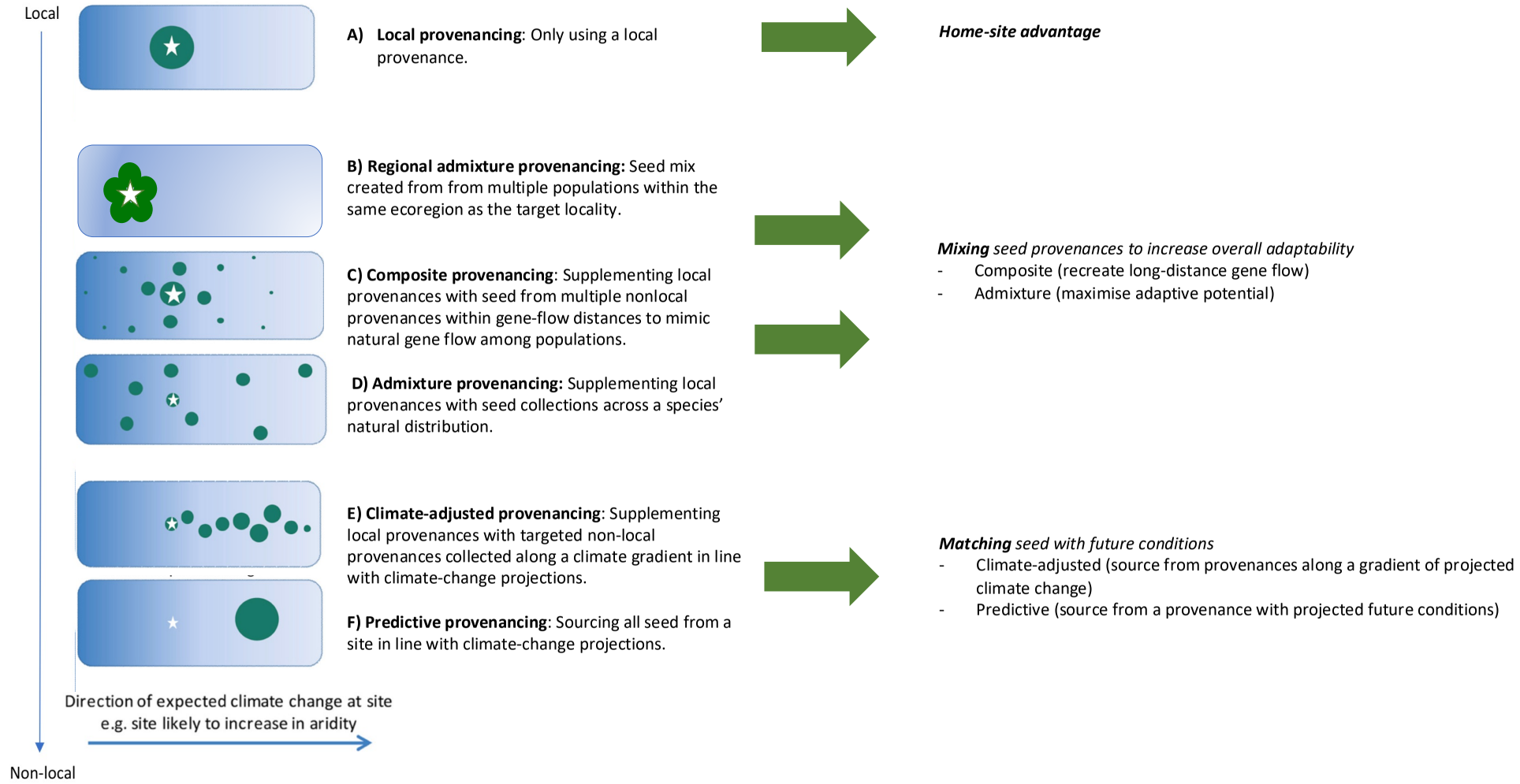


Figure 1. Main seed sourcing strategies for restoration (modified from Bucharova et al., 2019, Prober et al., 2015, and Byrne et al., 2013). The star represents revegetation site, and the circles represent seed source sites. Circle size indicates quantity of seed to be taken from site. Arrows represent direction of expected environmental change.

Seed supply

Revegetation efforts will fail without access to largescale and sustainable supplies of seed, regardless of whether a consensus is found over suitable provenancing methods, (Broadhurst et al., 2015, Mortlock, 2000a). Vast amounts of seed are required for such projects and meeting supply demands is a major challenge. For example, in a project to restore eelgrass (*Zostera marina*), 37.8 million viable seeds were added to over 125.2 ha area in a one-year time period (2009-2010) (Orth et al., 2012). With continued declines in the abundance and health of native plant populations due to habitat fragmentation and climate change, seed supplies are increasingly limited and seed source populations can be exposed to overharvesting. The result is a 'restoration paradox' – the precarious balance between accessing enough high-quality seed for restoration whilst not overharvesting wild populations (Broadhurst et al., 2015, Merritt and Dixon, 2011, Mortlock, 2000a).

Establishing seed production areas (SPAs; also known as seed orchards) may help to overcome seed supply issues (Broadhurst et al., 2016). SPAs are 'seed farms' created by planting the target restoration species to provide a sustainable seed supply for restoration. This can relieve pressure on native populations and create more reliable seed supplies, less dependent on unpredictable patterns of seeding and germplasm quality associated with wild-sourced seed (Broadhurst et al., 2015). SPAs allow practitioners greater control over the germplasm produced (e.g., parent plant health, provenance, breeding, genetic composition), and allow a more custom-fit approach to seed supply — i.e., breeding of populations able to cope with specific environmental issues such as soil toxicity (Broadhurst et al., 2016). Establishment of a SPA can, however, be complex, but addressing these four key questions will increase success (Nevill et al., 2016):

1. A sound understanding of target species biology — this could include flower/fruit phenology, growth requirements, and pollinator mutualisms.
2. SPA design – including catering to the habitat requirements of essential pollinators and using an appropriate planting type to maximise germplasm quality (e.g., location, mixed planting vs monoculture).
3. Genetic considerations – selecting suitable provenances and maintaining genetic integrity by avoiding genetic contamination through accidental breeding.
4. Economics and society – economic and social feasibility of the SPA and consideration of partnerships through community, government, industry, and research.

Restoration and conservation genetics

Most of the seed sourcing strategies described above, and indeed any genetic rescue programme, require significant genetic considerations of the species involved. There are four relevant fields of genetics in this space.

1. Quantitative genetics and the genetic basis of functional traits.
2. Ecological genetics of contemporary gene flow and mating systems.
3. Population genetic structure and estimating historical demography.
4. Adaptation genetics of functional or fitness genes.

These areas and their key research methods are outlined below, specifically in the context of seed sourcing for restoration and conservation programmes.

Quantitative genetics and the genetic basis of functional traits

A functional trait can be defined as "morphological, biochemical, physiological, structural, phenological, or behavioural characteristics that are expressed in phenotypes of individual organisms and are considered relevant to the response of such organisms to the environment and/or their effects on ecosystem

properties" (Violle et al. 2017). Primarily, traits determine how plants respond to their environment, but they also influence both the interspecific and intraspecific interactions between organisms and have a direct effect on ecosystem processes (Kattge et al., 2011). These functional traits can be loosely separated into ecophysiological traits and life-history traits (Violle et al., 2007). Typically, ecophysiological traits involve a physiological response to the environment, such as photosynthesis, transpiration, and leaf temperature. Established ecophysiological trait measurements include specific leaf area (SLA), stomatal size and density, and nitrogen content per unit mass. Life-history traits associate with dispersal in time and space, establishment, fecundity, growth rate, and phenology (Violle et al., 2007, Weiher, 2007). These trait measurements classically include seed mass, which correlates with dispersal distance, fecundity, and establishment; SLA, which is linked to relative growth rate (RGR); and measuring the onset of flowering which is indicative of phenology (Pérez-Harguindeguy et al., 2013, Weiher et al., 1999).

The range of functional values present in a population depends on both the genetic background of the population as well as the capacity of the plants to assimilate to their local environmental conditions via phenotypic plasticity (Broadhurst et al., 2017, Nicotra et. at., 2007, Whitham et al., 2006). In the context of revegetation, quantifying the variation in functional traits found in populations can be indicative of their response to certain environmental variables. For example, higher wood density has been directly linked to improved plant fitness in arid conditions (Hacke et al., 2001). Natural selection is the main driver of local adaptation and occurs as a result of interactions between the genetically determined phenotype and the environment. If local adaptation is present in a population, then the local genotype should have a higher relative fitness than those from other environments (Kawecki and Ebert, 2004) However, the homogenising effect of gene flow and a reduction in genetic diversity resulting from genetic drift can obstruct local adaptation via natural selection (Kawecki and Ebert, 2004).

Early quantitative genetic studies of local adaptation arose from interest in the variability present in the traits of plant species growing along environmental gradients such as altitude. One of the first trials exploring the genetics of plant populations focussed on *Potentilla tormentilla,* a relative of the strawberry found in the Swiss Alps. Clones were taken from parent plants growing in a median altitude and were cultivated in areas of the highest and lowest altitudes of the species range, allowing researchers to tease apart phenotypic plasticity from plant genetic responses (Bonnier, 1894, Fuller, 1921, Hiesey, 1940). Today, studies of local plant adaptation remain fundamentally similar to those early study mentioned above.

The two main types of experiment used are reciprocal transplant trials (transplanting individuals between habitats) and common garden trials (growing individuals from a variety of habitats under common conditions). These trials can be used to determine where local adaptation may be occurring, and which functional traits are affected. Such information can be utilised when designing seed zones which aim to avoid introducing plants with maladapted traits. These results could also be used in climate change mitigation strategies, such as climate predictive seed sourcing (Prober et al., 2015, Ramalho et al., 2017). Adaptation could be 'fast-tracked' by mixing genotypes from sites with functional traits adapted to the predicted future climate at the restoration site with the local provenance. Reciprocal transplant and common garden trials have been used on a large scale in forestry for decades, allowing practitioners to determine which populations would be most successful for plantation and therefore the most lucrative. Some of the processes utilised in this context are being readily applied to restoration (Gellie et al., 2016, Havens et al., 2015, Hodgins and Moore, 2016).

Reciprocal transplant trials are carried out in the field and involve transplanting individuals between habitats with differing environmental conditions. Then, functional traits can be measured to see if individuals perform best in their home-site conditions, which is indicative of local adaptation. Using a reciprocal

transplant design, Breed et al., (2016) explored the level of local adaptation in *Eucalyptus socialis* in two sites of contrasting aridity in South Australia. Seed was collected from both provenances and a total of 476 *E. socialis* seedlings were raised in glasshouses then divided equally and transplanted into randomised plots in the two locations. Height and survival rates were then monitored over five years and local adaptation was assessed for the two provenances. The only home-site advantage detected was an increase in height of the arid provenance grown at the arid site, suggesting adaptation to aridity within the species. In the context of seed sourcing for restoration, the study suggested that, as climate shifts to warmer temperatures, the mixing of seed sourced in the arid environment with local seed to be planted in the mesic habitat may help create climate change-resilient populations of *E. socialis* in the mesic habitat by mixing arid adapted seed with the local provenance. However, while reciprocal transplant trials are considered the best method to detect local adaptation (Gibson et al., 2016, Kawecki and Ebert, 2004), such trials are often impractical to conduct for logistical or legal reasons (Kawecki and Ebert, 2004). As such, common garden trials are easier to implement and more commonly used.

A common garden trial involves taking individuals from a variety of populations found in contrasting habitats and growing them in a controlled environment such as a greenhouse or an experimental plot. It is then possible to manipulate conditions to suit the experiment, either by imitating environmental conditions to emulate a reciprocal transplant trial, or with uniform conditions and thus distinguish genetic effects from phenotypic plasticity (Kawecki and Ebert, 2004). Using a common garden design, Hancock et al. (2013) looked for evidence of local adaptation in six plant species in New South Wales (*Acacia falcata, Bursaria spinosa* subsp. *spinosa, Eucalyptus crebra, E. tereticornis, Hardenbergia violacea* and *Themeda australis*), representative of a variety of growth forms. For each species, seed was collected from a range of geographic provenances with contrasting habitats, including one local to the experimental plots used in the study. Seeds were germinated and

raised in glasshouses under common conditions, and then planted into experimental plots. Growing time differed for each species (ranging from 7 months to 2 years), and functional traits relating to survivorship, growth, phenology, morphology, and herbivory were measured. Local adaptation was consistently detected only in one species (*B. spinosa*), and in the survivorship rates of *T. australis*. In the context of seed sourcing for restoration, the study recommended that, as there were generally low levels of local adaptation, using non-local seed could be considered in the region in order to increase the evolutionary potential in the face of changing environmental conditions.

The design of many existing reciprocal transplant and common garden studies may make them of limited use for informing restoration projects. Gibson et al. (2016) outlined six experimental considerations key to the successful application of local adaptation research to restoration practice, including the experimental environment, response variables, among-population variability, the selective agent, maternal effects, and spatial and temporal variability. Gibson et al. (2016) also outlined methodologies in experiments that help take into account these considerations, including using multiple sites in trials, incorporating a measurement of population growth, studying multiple generations, and testing for adaptation to biotic factors such as the native plant community. However, incorporating such methodologies can be time consuming and expensive, especially in long lived plants such as trees. The review by Gibson et al. (2016) also highlighted many inconsistencies with the use of these methodologies in the literature and suggested a degree of caution is required for practitioners wishing to implement these studies into seed sourcing decision making.

Mating systems

In hermaphroditic plant species, there are three primary mating systems — autogamy (self-pollination), outcrossing (fertilisation by other individuals), and mixed mating (a combination of the previous two) (Imai et al., 2016). For plants,

mating systems are a compromise — while autogamy provides greater reproductive assurance, it comes at the cost of avoiding sexual reproduction (Imai et al., 2016, Schoen et al., 1996). Outcrossing, on the other hand, is a key process in the introduction of new alleles to a population and the maintenance of genetic diversity via sexual reproduction (Hamrick and Godt, 1996). However, this mating system is limited by the availability of potential mates and, in the case of animal-pollinated angiosperms, the presence of suitable pollinators (Delmas et al., 2016, Eckert et al., 2010). There is evidence to suggest that habitat fragmentation can reduce the levels of outcrossing found in plant species (Aguilar et al., 2006, Aguilar et al., 2008, Breed et al., 2012, Eckert et al., 2010). Land disturbance can alter pollination through changing plant and pollinator densities, and pollinator behaviour (Hadley and Betts, 2012). Changes in these key mechanisms have led to a rise in levels of inbreeding and, as a result, autogamy being more commonly adopted as a mating system (Aguilar et al., 2006, Aguilar et al., 2008, Eckert et al., 2010). In a meta-analysis Aguilar et al. (2006) found that habitat fragmentation was associated with an increase in selfing in plant populations. Further, in a desktop review, Eckert et al. (2010) found progeny of fragmented populations to show both higher levels of inbreeding and lower levels of outcrossing in comparison to non-fragmented habitats. Changes in pollinator behaviour can impact plant mating systems. For example, if the next pollen source has a further distance to travel due to fragmentation, then pollinators are less likely to make the journey, instead staying on the same plant for longer, and increasing the necessity of inbreeding and selfing (Breed et al., 2015, Hadley and Betts, 2012). In a study of the impact of habitat fragmentation on the genetic diversity of three mallee eucalypt species (*Eucalyptus incrassata, E. gracilis*, and *E. socialis*), Breed et al. (2015) found that the predominantly bird pollinated species (*E. incrassata*) demonstrated lower levels of inbreeding compared to the two insect-pollinated species. As birds have far greater mobility than invertebrates, this could indicate that increased pollinator mobility may increase genetic diversity. Further, in a meta-analysis, Breed et al. (2015) showed habitat fragmentation to generally increase the likelihood of

plants to self-pollinate, but species with more mobile pollinators where somewhat less likely to experience this transition.

Ecological genetics of contemporary gene flow

Gene flow can be defined as "a collective term that includes all mechanisms resulting in the movement of genes from one population to another" (Slatkin, 1985), and occurs through the "movement of gametes, individuals and even entire populations" (Slatkin, 1987). While levels of genetic diversity have been shown to decrease as a result of genetic drift, habitat fragmentation, and natural selection, high levels of gene flow have been shown to reduce these effects (Lenormand, 2002). As such, gene flow can effectively homogenize population gene pools, counteracting local adaptation. In the context of seed sourcing for restoration, contemporary gene flow can have significant implications for seed quality – if fragmented populations demonstrate limited gene flow, then the consequence will be significantly lowered levels of genetic diversity and likely increased levels of inbreeding.

Contemporary gene flow assessments can determine whether landscape changes have compromised the evolutionary potential of a population and are achieved by assessing the flow of propagules such as pollen and seed between populations (Sork and Smouse, 2006). Typically, the distance and rate of gene flow is estimated directly after it has occurred through pollination. It is possible to estimate actual gene movement (e.g., via established seedlings), but the more commonly used method is to estimate potential gene movement (e.g., the fertilised seed on maternal plants) as an effect of pollen flow (Sork and Smouse, 2006, Sork et al. 1999). Gene flow assessments can then be used to ascertain the number of pollen sources, define the dispersal curve of the pollen, and detect the proportion of pollen which migrates from beyond the given study area (Sork and Smouse, 2006, Smouse et al., 2001). One method is a paternity assessment – defining the source and location of pollen which has fertilised the seeds on given maternal plants by conducting a paternity assignment

through genotyping (Sork, 2016, Sork and Smouse, 2006). Byrne et al. (2008) assessed gene flow in remnant populations of *Eucalyptus wandoo*, a predominantly invertebrate pollinated tree found in southern Western Australia. A paternity assignment test demonstrated that, despite fragmentation of the habitat of this species, there were still high levels of contemporary gene flow. This was attributed to isolated individuals (i.e., paddock trees) maintaining a genetic connectivity with fragmented populations. However, paternity assignments usually involve an exhaustive sampling of all individuals in the study area and can often be difficult to implement on a large scale. The TwoGener method (Smouse et al., 2001) is an alternative method which assesses pollen cloud structure through the composition of genes in the pollen pool and models contemporary gene flow. This removes the need for a paternity analysis, making the TwoGener a more feasible method on a landscape scale (Austerlitz et al., 2004, Austerlitz and Smouse, 2001, Sork, 2016, Sork and Smouse, 2006).

Population genetic structure and estimating historical demography

Neutral genetic variation (i.e., genetic variation that is not shaped by natural selection) within populations is shaped by a number of attributes including life history (e.g., growth form), demography (e.g., changes in abundance), and geography (e.g., range size) (Broadhurst et al., 2017, Hamrick and Godt, 1996). Interpretation of neutral genetic variation is utilised in restoration and conservation genomics for deciphering historical genetic structure and gene flow among populations (Broadhurst et al., 2017) in order to inform the relatedness and connectivity among contemporary populations. Analysis requires a dataset of neutral genetic markers, such as single nucleotide polymorphisms (SNPs) or microsatellites.

Several key population genetic statistics exist that are directly relevant to restoration and conservation. The most common measurement of neutral genetic variation is observed and expected mean heterozygosity (H_O, H_E), which estimate the heterozygosity of each individual at each locus (Frankham, 1989). Low levels of

heterozygosity within a population (i.e., many/most genetic markers being homozygous) can be indicative of inbreeding and/or genetic drift. Inbreeding depression, whereby the reduction in genetic diversity resulting from inbreeding results in reduced fitness, occurs most often after population bottlenecks (Reed and Frankham, 2003). High levels of genetic drift are generally experienced by isolated and/or fragmented populations (Breed et al., 2013). While neutral genetic diversity does not imply evolutionary potential (Holderegger et al., 2006), higher levels of neutral diversity have been linked to increased fitness (Leimu et al., 2006). While calculating H_O and H_E remains the most popular method of estimating genetic variation, this metric has been shown to be less sensitive in estimating recent or short-term changes to populations than alternative methods (Gonzalez et al., 2020; Allendorf, 1986). One such method is to quantify the allelic richness of a population through estimating the actual number of alleles present at each locus (Allendorf, 1986). A benefit of this method is that even the loss of rare alleles from a recent event (e.g., population bottleneck) can be detected, whereas simply scoring a locus as heterozygous or homozygous may miss these nuances (Gonzalez et al., 2020; Allendorf, 1986).

Measures of among-population genetic differentiation (F_{ST}) can inform the genetic structure of populations and the levels of genetic differentiation among them. Pairwise F_{ST} allows a comparison of the frequency of alleles between populations from a scale of zero (the populations share all alleles) to 1 (the populations share no alleles) and is indicative of both gene flow and population structure. There are other analytical methods to estimate the most likely number of genetic clusters (K) which can be used to visualise population structure (Figure 2.). Options include a Discriminant Analysis of Principal Components (DAPC) (Jombart et al., 2010), and STRUCTURE (Pritchard et al., 2000), and ADMIXTURE (Alexander et al., 2009).

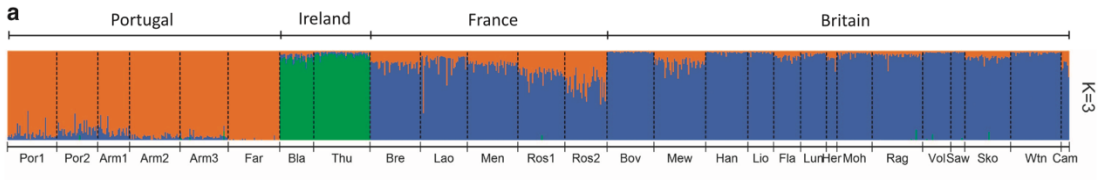


Figure 2. Example of a STRUCTURE analysis for *Eunicella verrucosa* (Holland et al., 2017). Three genetic clusters have been identified (K=3) and each corresponds with a colour. Each vertical bar represents an individual.

Patterns of population genetic structure across geographic areas may be shared among species if, for example, there are common barriers to gene flow for each species across the region. In this situation, this information can be used in the delineation of generalised seed zones. For example, Jørgensen et al. (2016) applied this method to delineate seed zones in Norway. Population structure was analysed for six alpine species commonly used in restoration projects with samples collected from across Norway, with results contradicting the existing recommendations of collecting local seed. Instead, a maximum of two genetic clusters per species was found and, in two instances, no genetic structure. The study recommended the relaxation of existing guidelines and suggested four general seed transfer zones for the whole country. There has been some attempt to link associations between genetic structure and attributes such as life history, range size and disjunction, and biome (Broadhurst et al., 2017, Ellegren and Galtier, 2016, Hamrick and Godt, 1996). Generalising links between population structure and attributes could theoretically assist seed source provenancing when no genetic data is available (i.e., if a species has a number of attributes that have been consistently linked with low genetic structure/high levels of gene flow, then relaxed provenancing can be prescribed) (Broadhurst et al., 2017). For example, Hamrick and Godt (1996) found breeding system and life form to be highly linked to population structure and genetic diversity. Outcrossing species displayed higher genetic diversity and lower population structure while short-lived annuals displayed lower genetic diversity and greater population structure. Caution must, however, be present in applying sweeping generalisations to restoration practice. A meta-analysis by Broadhurst et al. (2017) found that some of the associations between

attributes and genetic diversity in the Australian flora differed to global trends. For example, Australian herbs showed lower genetic differentiation compared to the global trends, despite the majority being invertebrate pollinated which is generally associated with higher population differentiation than plants pollinated by wind or birds.

Adaptive genetics of functional genes

Understanding the levels of genetic diversity in a population can be suggestive of its ability to adapt to a changing environment, as higher levels of adaptive genetic diversity provide a larger pool of functional genetic variation for selection to act upon (Jump and Peñuelas, 2005). By looking for associations between the genetic diversity of populations and their environment, researchers can identify putative genomic targets of environmental selection (e.g., Christmas et al., 2016). The adaptive relevance of such genomic regions can then be identified either through comparisons with a reference genome or, if unavailable, through the identification of orthologous sequences in related species via BLAST searches (Altschul et al., 1997). Finally, combining this with the environmental data now often freely available (e.g., Atlas of Living Australia), researchers have great scope to inexpensively explore how the environment is shaping the genome of a species of interest.

Outlier tests, such as those used by Bayescan (Foll and Gaggiotti, 2008), which look for F_{ST} outlier loci within a set of genetic markers (i.e., genetic markers that display a greater or lower level of genetic differentiation than expected under a neutral model) have been effectively employed by a number of studies to identify targets of selection among populations (Christmas et al., 2016, De Kort et al., 2014b, Steane et al., 2017). Gene-environment association analyses used by packages such as Bayenv and LFMM (Coop et al., 2010, Frichot et al., 2013) can be used to not only identify genomic targets of selection but also to identify the environmental selective agents. Christmas et al. (2016) explored genetic associations with environmental variables in the non-

model species *Dodonaea viscosa* subsp. *angustissima* to screen for adaptive variants. Single-nucleotide polymorphisms (SNPs) were analysed for 17 populations spanning a latitudinal gradient. A host of genes, with functions relating to water use efficiency as well as other environmental responses, were found to contain SNPs that correlated with water availability, temperature, and elevation. The study suggested the such findings can be applied in a restoration context, by informing climate-adjusted seed provenancing.

Genetic Marker Technologies

In a relatively short period of time, there has been a remarkable increase in the genetic tools available to genetic research for restoring and conserving plant species (Breed et al., 2019, Williams et al., 2014, Primmer 2009). Traditionally, allozymes were the main molecular markers used in genetic research but offered limited coverage of the genome as only a low number of loci are examined (Williams et al., 2014). The arrival of PCR in the 1990s allowed the development of molecular markers which can amplify and analyse specific regions of DNA (Eeles et al., 1992). Molecular markers that have emerged since the development of PCR include amplified fragment length polymorphisms (AFLP's), microsatellites, and the more recent development of single nucleotide polymorphisms (SNPs) which has become more commonplace since the development of next generation sequencing (NGS). NGS can sequence millions of small DNA fragments simultaneously, thus drastically reducing the time needed to sequence genomes or parts of many genomes in comparison to previous Sanger technology (Behjati and Tarpey, 2013). Traditional molecular markers were putatively neutral (i.e., AFLPs, allozymes, and microsatellites) (Narum et al., 2013, Reed and Frankham, 2003) making adaptive analyses difficult. The recent increase in sequencing capacity has allowed generation of genome-wide SNP markers, including SNPs in both adaptive and neutral gene regions (Narum et al., 2013). The development of genome-wide SNPs presents the opportunity to explore adaptive evolution in order to define seed zones (Sork, 2016). Further, it was recently emphasised that, for

improved accuracy and therefore confidence in seed zones based on genetic structure, a genome-wide sequencing approach is required (Shafer et al., 2015). Studies basing seed zone recommendations on just a few markers (e.g., microsatellites) or dominant markers (e.g., AFLP markers) are prone to known biases, as they only represent a small and hypervariable proportion of the genome or have well-established problems in estimating genetic structure (Ellegren, 2004, Mueller and Wolfenbarger, 1999). An example of an NGS method whith uses SNP data is restriction-site associated DNA sequencing (RAD-seq), which identifies markers randomly across the genome (Davey and Blaxter, 2011). While this method can produce datasets with many thousands of markers spanning adaptive and neutral regions of the genome, the information accessed is at a much reduced representation of the whole genome (Davey and Blaxter, 2011). Due to this, the effectiveness of RADseq for studies of local adaptation has been questioned, as sites under selection may be missed (Lowry et al. 2017). As such, the effectiveness of methods such as RAD-seq for adaptive studies requires exploration.

Application and implications

While quantitative genetic studies will continue to play a vital role in detecting local adaptation and measuring functional traits (de Villemereuil et al., 2016), the application of such methods is time consuming and expensive. The relative absence of quantitative genetic studies so far in restoration is partly attributable to this factor (Broadhurst et al., 2017). The arrival of next generation sequencing (NGS) has revolutionised biology in general (Schuster, 2008), but it also delivers the potential to apply genomics in restoration and conservation practice (Breed et al., 2019, Williams et al., 2014, Primmer 2009). Prior to NGS, generating genomic data was costly, time consuming, and therefore mostly impractical in the context of restoration. Furthermore, prior to the development of methods to rapidly and cheaply analyse SNPs (e.g., RADseq data) which provide access to regions of the genome under selection, most molecular markers were assumed to be putatively neutral, limiting

adaptive studies (Holderegger et al., 2006). As such, practitioners who wish to understand population structure and/or adaptive genetic variation in a species can now employ these technologies to achieve rapid results in a far more cost-effective way than by using traditional quantitative methods. However, despite an increase in publications in this area, there is still considerable scope to improve our understanding of the genetic structure and adaptation of plants used in restoration and conservation projects, particularly through the use of genomics (Breed et al., 2019, Williams et al., 2014, Primmer 2009). It is still therefore uncommon for plant species to have their seed sourcing methods defined by population genomics (but see Durka et al., 2017, Hufford et al., 2016, Jørgensen et al., 2016, De Kort et al., 2014a, Jørgensen et al., 2016, Krauss et al., 2013, Stingemore and Krauss, 2013).

In restoration practice, there is evidence to suggest an imminent shift away from the 'local is best' standpoint to a focus on the maintenance or improvement of genetic diversity in populations. For example, the recently published *National Standards for the Practice of Ecological Restoration in Australia* (NSPERA; McDonald et al., 2016) states that preserving the genetic diversity of populations should be the primary concern. NSPERA promote optimising gene flow through the connection of fragmented habitats and, when the local provenance appears inbred, the introduction of alleles from larger populations (even if not sourced locally). In addition, NSPERA outlined the potential of climate-adjusted provenancing (Prober et al., 2015, Ramalho et al., 2017), but an emphasis is placed on the need for more research in order to develop 'rules of thumb' in seed sourcing.

Knowledge gaps and research opportunities

The successful integration of new methods for seed sourcing techniques adopted by practitioners requires empirical studies. However, despite the huge investments that we are seeing in conservation and ecosystem restoration, genetic research is still lacking. For example, at the time of starting this thesis (2017), in South

Australia there were only three published studies that have looked at neutral or adaptive genetic diversity with relevance to seed zoning, and these studies focus only on two species (Christmas et al., 2017, Christmas et al., 2016, McCallum et al., 2013). If the management of plant species is to maintain the current trajectory of shifting towards genetic based decision making, then more research is integral. Under a continually shifting climate, seed sourcing recommendations are increasingly suggestive of facilitated gene flow to provide populations with adaptations which will help them cope with the changes to come (Breed et al., 2013, Broadhurst et al., 2008, Prober et al., 2015, Ramalho et al., 2017, Sgrò et al., 2011, Williams et al., 2014). Existing studies predominantly focus on population structure (Broadhurst et al., 2017), and while this is indicative of historical demography and a viable method for defining seed zones, it cannot tell which local adaptations of a species correlate with specific environmental variables (Sork, 2016). Indeed, neutral genetic diversity may show conflicting patterns to that of adaptive variation (Holderegger et al., 2006). If the future of seed sourcing for restoration is to incorporate assisted gene flow, then future research must incorporate adaptive genomics (Hodgins and Moore, 2016). While there remains a lack of neutral genetic studies in relation to seed sourcing for restoration, there are even fewer adaptive studies (but see Christmas et al., 2016, Hufford et al., 2016, Shryock et al., 2017, Steane et al., 2014). The recent arrival of NGS technology have made adaptive studies feasible and affordable (Williams et al., 2014). This development promises adaptive studies to have an exciting role in the future of ecological restoration (Sork, 2016).

Thesis overview

For this thesis, range-wide samples were collected for three native Australian plant species (all lacking a reference genome), and population-level reduced representation genome sequencing was generated for each. Each species has different life history traits (e.g., pollination syndrome, range, lifespan), and different suite of management concerns and priorities (e.g., genetic rescue, species boundaries, seed

zone delineation) commonplace in the management of plant species. While the management concerns may differ, addressing them requires many of the core tenets of conservation and restoration genetics to be resolved (e.g., genetic diversity and structure, genetic adaptation). The aim of this thesis is to generate and analyse genomic datasets to answer a wide variety of management questions for three plant species in southern Australia (see Figure 3. for the workflow to be applied for each chapter). These three studies test the use of genomics in addressing questions about genetic structure, diversity, and adaptation in order to develop management guidelines based on the principles of conservation and restoration genetics. By doing so, this also addresses one of the core challenges to the uptake of genomics into restoration and conservation programmes: the translation of results into practice (Shafer et al., 2015).

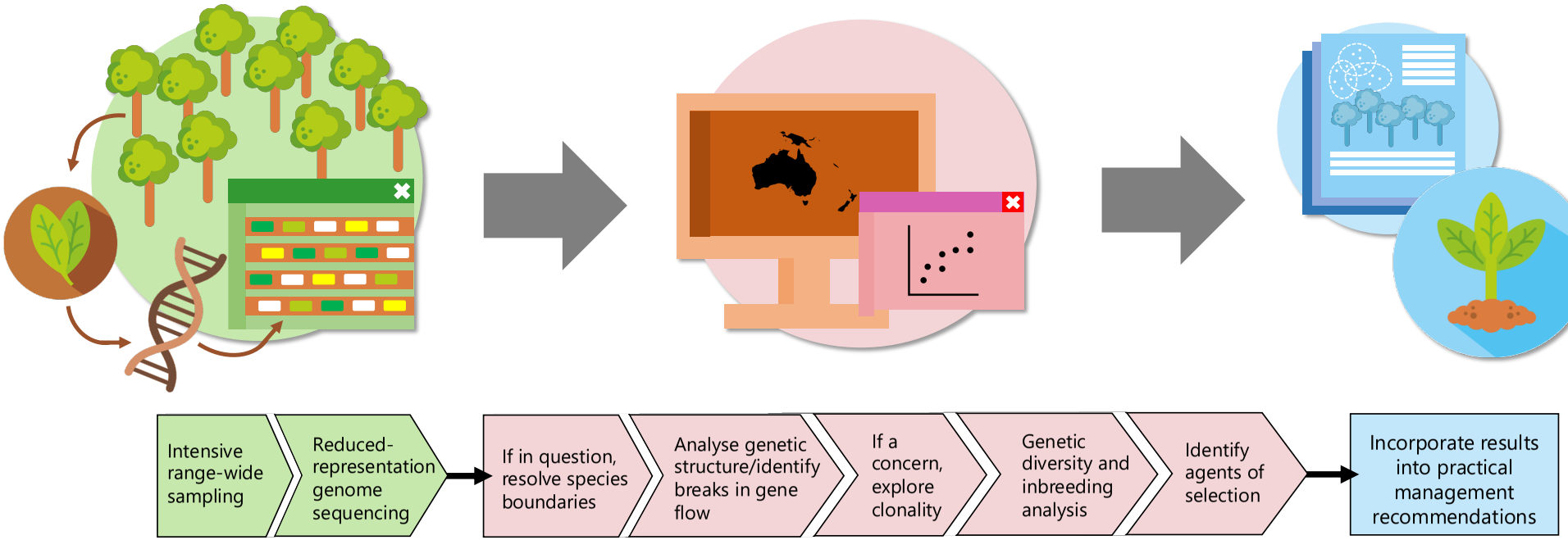


Figure 3. Planned workflow to be applied in each PhD chapter. Figure modified from the workflow for conservation genomic studies of threatened plant taxa by Rossetto et al. (2021).

Chapter 2. The conservation management of the endangered Whibley wattle (*Acacia whibleyana*). *A. whiblyeana* is one of Australia's most endangered wattle species with an extremely restricted and fragmented range. Due to low seed set and recruitment in the species, a priority management solution is to establish translocated populations. This chapter explores the genetic diversity and structure of all known remnant stands, a translocated population, and a post-fire seedling cohort of this translocation, to explore variation of genetic diversity and structure. Results are used to give direct management recommendations.

Chapter 3. Revegetation strategies for the widespread and abundant pearl bluebush (*Maireana sedifolia*), which is heavily used in post-mining rehabilitation. *M. sedifolia*, a long-lived, keystone arid zone shrub with a large, continuous range. It is an essential species in largescale arid zone revegetation projects, yet there are several challenges over acquiring enough seed. For this chapter, a range-wide reduced representation dataset is used to explore variation of genetic diversity and structure to delineate seed sourcing guidelines for the species. Links between genetic variation and environmental selection pressures are also explored alongside a comparison of the environments of the genetic clusters found in the study.

Chapter 4. Species boundaries and conservation management of the wide-ranging, yet vulnerable daisy bush *Olearia pannosa*. The species is found across southeastern Australia in fragmented and scattered populations. There are currently two subspecies (*O. pannosa.* subsp. *pannosa* and *O. pannosa* subsp. *cardiophylla*), differentiated by leaf morphology, but some clarification is necessary over the genetic divergence between them. This chapter uses genomics firstly to explore and clarify the subspecies boundaries. Then, the genetic diversity and divergence with and between populations is explored and compared between the two subspecies and the likely new species identified in the genetic divergence analysis. For *O. pannosa.* subsp. *pannosa*, the links between genetic variation and environmental selection pressures are also explored**.**

Chapter 5. A general discussion of the thesis, summarising the outcomes of each chapter and discussing them in the context of whether each of the components of the planned workflow (Figure 3) were achieved.

References

Aguilar, R., Ashworth, L., Galetto, L. & Aizen, M. A. 2006. Plant reproductive susceptibility to habitat fragmentation: review and synthesis through a meta-analysis. *Ecology Letters,* 9**,** 968-980.

Aguilar, R., Quesada, M., Ashworth, L., Herrerias-Diego, Y. & Lobo, J. 2008. Genetic consequences of habitat fragmentation in plant populations: susceptible signals in plant traits and methodological approaches. *Molecular Ecology,* 17**,** 5177-5188.

Aitken, S. N. & Whitlock, M. C. 2013. Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics,* 44**,** 367.

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T. & Curtis-Mclane, S. 2008. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications,* 1**,** 95-111.

Alexander, D. H., Novembre, J. & Lange, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome research,* 19**,** 1655-1664.

Allendorf, F.W., 1986. Genetic drift and the loss of alleles versus heterozygosity. *Zoo biology*, *5*(2), pp.181-190.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research,* 25**,** 3389-3402.

Austerlitz, F., Dick, C. W., Dutech, C., Klein, E. K., Oddou-Muratorio, S., Smouse, P. E. & Sork, V. L. 2004. Using genetic markers to estimate the pollen dispersal curve. *Molecular Ecology,* 13**,** 937-954.

Austerlitz, F. & Smouse, P. E. 2001. Two-generation analysis of pollen flow across a landscape. II. Relation between (Phi)ft, pollen dispersal and interfemale distance. *Genetics,* 157**,** 851-857.

Bai, Z. G., Dent, D. L., Olsson, L. & Schaepman, M. E. 2008. Proxy global assessment of land degradation. *Soil use and management,* 24**,** 223-234.

British Broadcasting Corporation (BBC). 2016. Community group buys historic Loch Arkaig Forest. *BBC News*.

Behjati, S. & Tarpey, P. S. 2013. What is next generation sequencing? *Archives of Disease in Childhood-Education and Practice,* 98**,** 236-238.

Bonnier, G. 1894. Les plantes arctiques comparees aux memes especes des Alpes et des Pyrenees. *Revue générale de botanique,* 6**,** 505-27.

Breed, M., Gellie, N. & Lowe, A. 2016. Height differences in two eucalypt provenances with contrasting levels of aridity. *Restoration Ecology,* 24**,** 471-478.

Breed, M. F. 2013. *Restoration Genetics of Murray Mallee and Neotropical Forests.* PhD thesis. University of Adelaide, Australia.

Breed, M. F., Gardner, M. G., Ottewell, K., Navarro, C. & Lowe, A. J. 2012. Shifts in reproductive assurance strategies and inbreeding costs associated with habitat fragmentation in Central American mahogany *Ecology Letters,* 15**,** 444-452.

Breed, M. F., Ottewell, K. M., Gardner, M. G., Marklund, M. H. K., Dormontt, E. D. & Lowe, A. J. 2015. Mating patterns and pollinator mobility are critical traits in forest fragmentation genetics. *Heredity,* 115**,** 108-114.

Breed, M. F., Stead, M. G., Ottewell, K. M., Gardner, M. G. & Lowe, A. J. 2013. Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. *Conservation Genetics,* 14**,** 1-10.

Broadhurst, L., Breed, M. F., Lowe, A., Bragg, J., Catullo, R., Coates, D., Encinas-Viso, F., Gellie, N., James, E., Krauss, S. & Byrne, M. 2017. Genetic diversity and structure of the Australian flora. *Diversity and Distributions,* 23**,** 41-52.

Broadhurst, L., Driver, M., Guja, L., North, T., Vanzella, B., Fifield, G., Bruce, S., Taylor, D. & Bush, D. 2015. Seeding the future – the issues of supply and demand in restoration in Australia. *Ecological Management & Restoration,* 16**,** 29-32.

Broadhurst, L. M., Jones, T. A., Smith, F. S., North, T. & Guja, L. 2016. Maximizing seed resources for restoration in an uncertain future. *BioScience,* 66**,** 73-79.

Broadhurst, L. M., Lowe, A., Coates, D. J., Cunningham, S. A., Mcdonald, M., Vesk, P. A. & Yates, C. 2008. Seed supply for broadscale restoration: maximizing evolutionary potential. *Evolutionary Applications,* 1**,** 587-597.

Bucharova, A., Bossdorf, O., Hölzel, N., Kollmann, J., Prasse, R. & Durka, W. 2019. Mix and match: regional admixture provenancing strikes a balance among different seed-sourcing strategies for ecological restoration. *Conservation Genetics,* 20**,** 7-17.

Byrne, M., Elliott, C., Yates, C. & Coates, D. 2008. Maintenance of high pollen dispersal in *Eucalyptus wandoo*, a dominant tree of the fragmented agricultural region in Western Australia. *Conservation Genetics,* 9**,** 97-105.

Byrne, M., Prober, S. M., McLean, E., Steane, D., Stock, W., Potts, B., et al. (2013). *Adaptation to Climate in Widespread Eucalypt Species*. Gold Coast, QLD: National Climate Change Adaptation Research Facility.

Christmas, M., Biffin, E., Breed, M. & Lowe, A. 2017. Targeted capture to assess neutral genomic variation in the narrow-leaf hopbush across a continental biodiversity refugium. *Scientific Reports,* 7**,** Publication number 41367.

Christmas, M. J., Biffin, E., Breed, M. F. & Lowe, A. J. 2016. Finding needles in a genomic haystack: Targeted capture identifies clear signatures of selection in a nonmodel plant species. *Molecular ecology,* 25**,** 4216-4233.

Christmas, M. J., Breed, M. F. & Lowe, A. J. 2015. Constraints to and conservation implications for climate change adaptation in plants. *Conservation Genetics*.

Clewell, A. & Rieger, J. P. 1997. What practitioners need from restoration ecologists. *Restoration Ecology,* 5**,** 350-354.

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P. & Skuce, A. 2013. Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental research letters,* 8**,** 024024.

Coop, G., Witonsky, D., Di Rienzo, A. & Pritchard, J. K. 2010. Using environmental correlations to identify loci underlying local adaptation. *Genetics,* 185**,** 1411-1423.

Costanza, R., D'arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'neill, R. V. & Paruelo, J. 1997. The value of the world's ecosystem services and natural capital. *nature,* 387**,** 253-260.

Crouzeilles, R., Curran, M., Ferreira, M. S., Lindenmayer, D. B., Grelle, C. E. & Benayas, J. M. R. 2016. A global meta-analysis on the ecological drivers of forest restoration success. *Nature Communications,* 7.

Davey, J.W. and Blaxter, M.L., 2010. RADSeq: next-generation population genetics. *Briefings in functional genomics*, *9*(5-6), pp.416-423.

De Kort, H., Mergeay, J., Vander Mijnsbrugge, K., Decocq, G., Maccherini, S., Kehlet Bruun, H. H., Honnay, O. & Vandepitte, K. 2014a. An evaluation of seed zone delineation using phenotypic and population genomic data on black alder Alnus glutinosa. *Journal of applied ecology,* 51**,** 1218-1227.

De Kort, H., Vandepitte, K., Bruun, H. H., Closset-Kopp, D., Honnay, O. & Mergeay, J. 2014b. Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species Alnus glutinosa. *Molecular Ecology,* 23**,** 4709-4721.

De Villemereuil, P., Gaggiotti, O. E., Mouterde, M. & Till-Bottraud, I. 2016. Common garden experiments in the genomic era: new perspectives and opportunities. *Heredity,* 116**,** 249.

Delmas, C. E., Fort, T. L., Escaravage, N. & Pornon, A. 2016. Pollen transfer in fragmented plant populations: insight from the pollen loads of pollinators and stigmas in a mass-flowering species. *Ecology and evolution,* 6**,** 5663-5673.

Durka, W., Michalski, S. G., Berendzen, K. W., Bossdorf, O., Bucharova, A., Hermann, J. M., Hölzel, N. & Kollmann, J. 2017. Genetic differentiation within multiple common grassland plants supports seed transfer zones for ecological restoration. *Journal of Applied Ecology,* 54**,** 116-126.

Eckert, C. G., Kalisz, S., Geber, M. A., Sargent, R., Elle, E., Cheptou, P.-O., Goodwillie, C., Johnston, M. O., Kelly, J. K., Moeller, D. A., Porcher, E., Ree, R. H., Vallejo-Marin, M. & Winn, A. A. 2010. Plant mating systems in a changing world. *Trends in Ecology & Evolution,* 25**,** 35-43.

Eeles, R. A., Warren, W. & Stamps, A. 1992. The PCR revolution. *European Journal Of Cancer,* 28**,** 289-293.

Ellegren, H. 2004. Microsatellites: simple sequences with complex evolution. *Nature reviews genetics,* 5**,** 435-445.

Ellegren, H. and Galtier, N., 2016. Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), pp.422-433.

Faast, R., Blyth, C., Breed, M., and Packer, J.G. 2019. *Acacia whibleyana*: Part 1 - Literature Synthesis to Inform Conservation of a Threatened Acacia. Report prepared by The University of Adelaide, on behalf of Department for Environment and Water, and Eyre Peninsula Natural Resources Management Board.

Fargione, J., Hill, J., Tilman, D., Polasky, S. & Hawthorne, P. 2008. Land clearing and the biofuel carbon debt. *Science,* 319**,** 1235-1238.

Feng, X., Fu, B., Lu, N., Zeng, Y. & Wu, B. 2013. How ecological restoration alters ecosystem services: an analysis of carbon sequestration in China's Loess Plateau. *Scientific reports,* 3**,** 2846.

Fisher, B. & Turner, R. K. 2008. Ecosystem services: classification for valuation. *Biological conservation,* 141**,** 1167-1169.

Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L. and Holling, C.S., 2004. Regime shifts, resilience, and biodiversity in ecosystem management. *Annu. Rev. Ecol. Evol. Syst.*, *35*, pp.557-581.

Foll, M. & Gaggiotti, O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics,* 180**,** 977-993.

Frankham, R. 2005. Genetics and extinction. *Biological Conservation,* 126**,** 131-140.

Frichot, E., Schoville, S. D., Bouchard, G. & François, O. 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular biology and evolution,* 30**,** 1687-1699.

Fuller, G. D. 1921. Alpine Adaptations. *Botanical Gazette,* 71**,** 405-406.

González, A.V., Gómez-Silva, V., Ramírez, M.J. and Fontúrbel, F.E., 2020. Meta-analysis of the differential effects of habitat fragmentation and degradation on plant genetic diversity. *Conservation Biology*, *34*(3), pp.711-720.

Gellie, N. J. C., Breed, M. F., Thurgate, N., Kennedy, S. A. & Lowe, A. J. 2016. Local maladaptation in a foundation tree species: Implications for restoration. *Biological Conservation,* 203**,** 226-232.

Gibbs, H. K. & Salmon, J. M. 2015. Mapping the world's degraded lands. *Applied Geography,* 57**,** 12-21.

Gibson, A. L., Espeland, E. K., Wagner, V. & Nelson, C. R. 2016. Can local adaptation research in plants inform selection of native plant materials? An analysis of experimental methodologies. *Evolutionary applications,* 9**,** 1219-1228.

Godefroid, S., Piazza, C., Rossi, G., Buord, S., Stevens, A.-D., Aguraiuja, R., Cowell, C., Weekley, C. W., Vogg, G., Iriondo, J. M., Johnson, I., Dixon, B., Gordon, D., Magnanon, S., Valentin, B., Bjureke, K., Koopman, R., Vicens, M., Virevaire, M. & Vanderborght, T. 2011. How successful are plant species reintroductions? *Biological Conservation,* 144**,** 672-682.

Hacke, U. G., Sperry, J. S., Pockman, W. T., Davis, S. D. & Mcculloh, K. A. 2001. Trends in wood density and structure are linked to prevention of xylem implosion by negative pressure. *Oecologia,* 126**,** 457-461.

Hadley, A. S. & Betts, M. G. 2012. The effects of landscape fragmentation on pollination dynamics: absence of evidence not evidence of absence. *Biological Reviews (Cambridge),* 87**,** 526-544.

Hamrick, J. L. & Godt, M. J. W. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society B-Biological Sciences,* 351**,** 1291-1298.

Hancock, N., Leishman, M. R. & Hughes, L. 2013. Testing the "local provenance" paradigm: a common garden experiment in Cumberland Plain Woodland, Sydney, Australia. *Restoration Ecology,* 21**,** 569-577.

Havens, K., Vitt, P., Still, S., Kramer, A. T., Fant, J. B. & Schatz, K. 2015. Seed sourcing for restoration in an era of climate change. *Natural Areas Journal,* 35**,** 122-133.

Hereford, J. 2009. A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist,* 173**,** 579-588.

Hiesey, W. M. 1940. Environmental influence and transplant experiments. *The Botanical Review,* 6**,** 181-203.

Hodgins, K. A. & Moore, J. L. 2016. Adapting to a warming world: ecological restoration, climate change, and genomics. *American journal of botany,* 103**,** 590-592.

Holderegger, R., Kamm, U. & Gugerli, F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology,* 21**,** 797-807.

Hoffmann, B.D. and Broadhurst, L.M., 2016. The economic cost of managing invasive species in Australia. *NeoBiota*, *31*, p.1.

Holland, L.P., Jenkins, T.L. and Stevens, J.R., 2017. Contrasting patterns of population structure and gene flow facilitate exploration of connectivity in two widely distributed temperate octocorals. *Heredity*, *119*(1), pp.35-48.

Honnay, O. & Jacquemyn, H. 2006. Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology,* 21**,** 823-831.

Hooke, R. L., Martín-Duque, J. F. & Pedraza, J. 2012. Land transformation by humans: a review. *GSA today,* 22**,** 4-10.

Hufford, K. M. & Mazer, S. J. 2003. Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution,* 18**,** 147-155.

Hufford, K. M., Veneklaas, E. J., Lambers, H. & Krauss, S. L. 2016. Genetic delineation of local provenance defines seed collection zones along a climate gradient. *AoB Plants***,** 8.

Imai, R., Tsuda, Y., Matsumoto, S., Ebihara, A. & Watano, Y. 2016. The Relationship between Mating System and Genetic Diversity in Diploid Sexual Populations of Cyrtomium falcatum in Japan. *PloS one,* 11**,** e0163683.

Jombart, T., Devillard, S. & Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics,* 11**,** 94.

Jørgensen, M. H., Elameen, A., Hofman, N., Klemsdal, S., Malaval, S. & Fjellheim, S. 2016. What's the meaning of local? Using molecular markers to define seed transfer zones for ecological restoration in Norway. *Evolutionary applications,* 9**,** 673-684.

Jump, A. & Peñuelas, J. 2005. Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters,* 8**,** 1010-1020.

Jusaitis, M. and Sorensen, B., 2007. Successful augmentation of an Acacia whibleyana (Whibley Wattle) population by translocation. *Australasian Plant Conservation: Journal of the Australian Network for Plant Conservation*, *16*(1), pp.23-24.

Kattge, J., Diaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P. B. & Wright, I. J. 2011. TRY–a global database of plant traits. *Global change biology,* 17**,** 2905-2935.

Kawecki, T. J. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecology Letters,* 7**,** 1225-1241.

Kennedy, C. M., Lonsdorf, E., Neel, M. C., Williams, N. M., Ricketts, T. H., Winfree, R., Bommarco, R., Brittain, C., Burley, A. L., Cariveau, D., Carvalheiro, L. G., Chacoff, N. P., Cunningham, S. A., Danforth, B. N., Dudenhöffer, J.-H., Elle, E., Gaines, H. R., Garibaldi, L. A., Gratton, C., Holzschuh, A., Isaacs, R., Javorek, S. K., Jha, S., Klein, A. M., Krewenka, K., Mandelik, Y., Mayfield, M. M., Morandin, L., Neame, L. A., Otieno, M., Park, M., Potts, S. G., Rundlöf, M., Saez, A., Steffan-Dewenter, I., Taki, H., Viana, B. F., Westphal, C., Wilson, J. K., Greenleaf, S. S.

& Kremen, C. 2013. A global quantitative synthesis of local and landscape effects on wild bee pollinators in agroecosystems. *Ecology Letters,* 16**,** 584-599.

Krauss, S. L., Sinclair, E. A., Bussell, J. D. & Hobbs, R. J. 2013. An ecological genetic delineation of local seed-source provenance for ecological restoration. *Ecology and evolution,* 3**,** 2138-2149.

Leimu, R. & Fischer, M. 2008. A meta-analysis of local adaptation in plants. *PLoS ONE,* 3**,** e4010.

Leimu, R., Mutikainen, P., Koricheva, J. & Fischer, M. 2006. How general are positive relationships between plant population size, fitness and genetic variation? *Journal of Ecology,* 94**,** 942-952.

Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution,* 17**,** 183-189.

Lesica, P. & Allendorf, F. W. 1999. Ecological genetics and the restoration of plant communities: mix or match? *Restoration Ecology,* 7**,** 42-50.

Lowe, A. J., Cavers, S., Boshier, D., Breed, M. F. & Hollingsworth, P. M. 2015. The resilience of forest fragmentation genetics - no longer a paradox - we were just looking in the wrong place. *Heredity,* 115**,** 97-99.

Lowry, D.B., Hoban, S., Kelley, J.L., Lotterhos, K.E., Reed, L.K., Antolin, M.F. and Storfer, A., 2017. Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation.

Maschinski, J. & Haskins, K. E. 2012. *Plant reintroduction in a changing climate: promises and perils,* Saint Louis, Missouri, Island Press.

Mccallum, K. P., Guerin, G. R., Breed, M. F. & Lowe, A. J. 2013. Combining population genetics, species distribution modelling and field assessments to understand a species vulnerability to climate change. *Austral Ecology***,** n/a-n/a.

Mcdonald, T., Jonson, J. & Dixon, K. W. 2016. National standards for the practice of ecological restoration in Australia. *Restoration Ecology,* 24**,** S4-S32.

Mcgill, B. J., Enquist, B. J., Weiher, E. & Westoby, M. 2006. Rebuilding community ecology from functional traits. *Trends in ecology & evolution,* 21**,** 178-185.

Mckay, J. K., Christian, C. E., Harrison, S. & Rice, K. J. 2005. "How local is local?" - a review of practical and conceptual issues in the genetics of restoration. *Restoration Ecology,* 13**,** 432-440.

Merritt, D. J. & Dixon, K. W. 2011. Restoration seed banks - a matter of scale. *Science,* 332**,** 424-425.

Mijnsbrugge, K. V., Bischoff, A. & Smith, B. 2010. A question of origin: Where and how to collect seed for ecological restoration. *Basic and Applied Ecology,* 11**,** 300-311.

Minnemeyer, S., Laestadius, L., Sizer, N., Saint Laurent, C. & Potapov, P. 2011. Bonn Challenge: a world of opportunity

Montalvo, A. M. & Ellstrand, N. C. 2000. Transplantation of the subshrub *Lotus scoparius*: testing the home-site advantage hypothesis. *Conservation Biology,* 14**,** 1034-1045.

Mortlock, W. 2000a. Local seed for revegetation: where will all that seed come from? *Ecological Management & Restoration,* 1**,** 93-101.

Mortlock, W. 2000b. Seed collection from woody plants for local revegetation. *FloraBank Guidelines.*

Mueller, U. G. & Wolfenbarger, L. L. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution,* 14**,** 389-394.

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R. & Hohenlohe, P. A. 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Molecular ecology,* 22**,** 2841-2847.

Nevill, P. G., Tomlinson, S., Elliott, C. P., Espeland, E. K., Dixon, K. W. & Merritt, D. J. 2016. Seed production areas for the global restoration challenge. *Ecology and Evolution,* 6**,** 7490-7497.

Nicotra, A.B., Atkin, O.K., Bonser, S.P., Davidson, A.M., Finnegan, E.J., Mathesius, U., Poot, P., Purugganan, M.D., Richards, C.L., Valladares, F. and van Kleunen, M., 2010. Plant phenotypic plasticity in a changing climate. *Trends in plant science, 15*(12), pp.684-692.

Nkonya, E., Mirzabaev, A. & Von Braun, J. 2016. Economics of land degradation and improvement: An introduction and overview. *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development.* Springer.

Oostermeijer, J.G.B., 2003. Threats to rare plant persistence. In *Population viability in plants* (pp. 17-58). Springer, Berlin, Heidelberg.

Orth, R. J., Moore, K. A., Marion, S. R., Wilcox, D. J. & Parrish, D. B. 2012. Seed addition facilitates eelgrass recovery in a coastal bay system. *Marine Ecology Progress Series,* 448**,** 177-195.

Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., Dasgupta, P. and Dubash, N.K., 2014. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change* (p. 151). IPCC.

Parmesan, C. 2006. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics,* 37**,** 637-669.

Pérez-Harguindeguy, N., Díaz, S., Garnier, E., Lavorel, S., Poorter, H., Jaureguiberry, P., Bret-Harte, M. S., Cornwell, W. K., Craine, J. M., Gurvich, D. E., Urcelay, C., Veneklaas, E. J., Reich, P. B., Poorter, L., Wright, I. J., Ray, P., Enrico, L., Pausas, J. G., De Vos, A. C., Buchmann, N., Funes, G., Quétier, F., Hodgson, J. G., Thompson, K., Morgan, H. D., Ter Steege, H., Van Der Heijden, M. G. A., Sack, L., Blonder, B., Poschlod, P., Vaieretti, M. V., Conti, G., Staver, A. C., Aquino, S. & Cornelissen, J. H. C. 2013. New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany,* 61**,** 167-234.

Pritchard, J., Stephens, M. & Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics,* 155**,** 945–959.

Primmer, C.R., 2009. From conservation genetics to conservation genomics. *Annals of the New York Academy of Sciences*, *1162*(1), pp.357-368.

Prober, S. M., Byrne, M., Mclean, E. H., Steane, D. A., Potts, B. M., Vaillancourt, R. E. & Stock, W. D. 2015. Climate-adjusted provenancing: a strategy for climate-resilient ecological restoration. *Frontiers in Ecology and Evolution,* 3.

Ramalho, C. E., Byrne, M. & Yates, C. J. 2017. A Climate-Oriented Approach to Support Decision-Making for Seed Provenance in Ecological Restoration. *Frontiers in Ecology and Evolution,* 5**,** 95.

Raven, P.H. and Yeates, D.K., 2007. Australian biodiversity: threats for the present, opportunities for the future. *Australian Journal of Entomology*, 46(3), pp.177-187.

Reed, D. H. & Frankham, R. 2003. Correlation between fitness and genetic diversity. *Conservation biology,* 17**,** 230-237.

Rossetto, M., Yap, J.Y.S., Lemmon, J., Bain, D., Bragg, J., Hogbin, P., Gallagher, R., Rutherford, S., Summerell, B. and Wilson, T.C., 2021. A conservation genomics workflow to guide practical management actions. *Global Ecology and Conservation*, *26*, p.e01492.

Rice, K. J. & Emery, N. C. 2003. Managing microevolution: restoration in the face of global change. *Frontiers in Ecology and the Environment,* 1**,** 469-478.

Sanderson, E. W., Jaiteh, M., Levy, M. A., Redford, K. H., Wannebo, A. V. & Woolmer, G. 2002. The human footprint and the last of the wild. *BioScience,* 52**,** 891-904.

Scheffers, B. R., De Meester, L., Bridge, T. C. L., Hoffmann, A. A., Pandolfi, J. M., Corlett, R. T., Butchart, S. H. M., Pearce-Kelly, P., Kovacs, K. M., Dudgeon, D., Pacifici, M., Rondinini, C., Foden, W. B., Martin, T. G., Mora, C., Bickford, D. & Watson, J. E. M. 2016. The broad footprint of climate change from genes to biomes to people. *Science,* 354.

Schoen, D. J., Morgan, M. T. & Bataillon, T. 1996. How does self-pollination evolve? Inferences from floral ecology and molecular genetic variation. *Philosophical Transactions: Biological Sciences***,** 1281-1290.

Schuster, S. C. 2008. Next-generation sequencing transforms today's biology. *Nature methods,* 5**,** 16.

Ser 2004. The SER international primer on ecological restoration, Version 2. Tucson, Arizona.: Society for Ecological Restoration Science and Policy Working Group.

Sgrò, C. M., Lowe, A. J. & Hoffmann, A. A. 2011. Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications,* 4**,** 326-337.

Shackelford, N., Hobbs, R. J., Burgar, J. M., Erickson, T. E., Fontaine, J. B., Laliberté, E., Ramalho, C. E., Perring, M. P. & Standish, R. J. 2013. Primed for change: developing ecological restoration for the 21st century. *Restoration Ecology,* 21**,** 297-304.

Shafer, A. B., Wolf, J. B., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., Colling, G., Dalen, L., De Meester, L. & Ekblom, R. 2015. Genomics and the challenging translation into conservation practice. *Trends in ecology & evolution,* 30**,** 78-87.

Shryock, D. F., Havrilla, C. A., Defalco, L. A., Esque, T. C., Custer, N. A. & Wood, T. E. 2017. Landscape genetic approaches to guide native plant restoration in the Mojave Desert. *Ecological Applications,* 27**,** 429-445.

Slatkin, M. 1985. Gene flow in natural populations. *Annual Review of Ecology and Systematics,* 16**,** 393-430.

Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science,* 236**,** 787-792.

Smouse, P.E., Dyer, R.J., Westfall, R.D. and Sork, V.L., 2001. Two-generation analysis of pollen flow across a landscape. I. Male gamete heterogeneity among females. Evolution, 55(2), pp.260-271.

Sork, V. L. 2016. Gene flow and natural selection shape spatial patterns of genes in tree populations: implications for evolutionary processes and applications. *Evolutionary applications,* 9**,** 291-310.

Sork, V. L. & Smouse, P. E. 2006. Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology,* 21**,** 821-836.

Sork, V.L., Nason, J., Campbell, D.R. and Fernandez, J.F., 1999. Landscape approaches to historical and contemporary gene flow in plants. *Trends in Ecology & Evolution*, *14*(6), pp.219-224.

Steane, D. A., Potts, B. M., Mclean, E., Prober, S. M., Stock, W. D., Vaillancourt, R. E. & Byrne, M. 2014. Genome-wide scans detect adaptation to aridity in a widespread forest tree species. *Molecular Ecology,* 23**,** 2500-2513.

Steane, D. A., Potts, B. M., Mclean, E. H., Collins, L., Holland, B. R., Prober, S. M., Stock, W. D., Vaillancourt, R. E. & Byrne, M. 2017. Genomic scans across three eucalypts suggest that adaptation to aridity is a genome-wide phenomenon. *Genome biology and evolution,* 9**,** 253-265.

Stingemore, J. A. & Krauss, S. L. 2013. Genetic delineation of local provenance in Persoonia longifolia: implications for seed sourcing for ecological restoration. *Restoration Ecology,* 21**,** 49-57.

Taylor, R.S., Watson, S.J., Bennett, A.F. and Clarke, M.F., 2013. Which fire management strategies benefit biodiversity? A landscape-perspective case study using birds in mallee ecosystems of south-eastern Australia. *Biological Conservation*, 159, pp.248-256.

Turesson, G. 1922. The genotypical response of the plant species to the habitat. *Hereditas,* 3**,** 211-350.

Verdone, M. & Seidl, A. 2017. Time, space, place, and the Bonn Challenge global forest restoration target. *Restoration Ecology*.

Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. 2007. Let the concept of trait be functional! *Oikos,* 116**,** 882-892.

Vranckx, G. U. Y., Jacquemyn, H., Muys, B. & Honnay, O. 2011. Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology,* 26**,** 228-237.

Weiher, E. 2007. On the status of restoration science: obstacles and opportunities. *Restoration Ecology,* 15**,** 340-343.

Weiher, E., Werf, A., Thompson, K., Roderick, M., Garnier, E. & Eriksson, O. 1999. Challenging Theophrastus: a common core list of plant traits for functional ecology. *Journal of vegetation science,* 10**,** 609-620.

Wilkinson, D. M. 2001. Is local provenance important in habitat creation? *Journal of Applied Ecology,* 38**,** 1371-1373.

Williams, A. V., Nevill, P. G. & Krauss, S. L. 2014. Next generation restoration genetics: applications and opportunities. *Trends in Plant Science,* 19**,** 529-537.

Whitham, T.G., Bailey, J.K., Schweitzer, J.A., Shuster, S.M., Bangert, R.K., LeRoy, C.J., Lonsdorf, E.V., Allan, G.J., DiFazio, S.P., Potts, B.M. and Fischer, D.G., 2006. A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, *7*(7), pp.510-523.

World Wide Fund for Nature (WWF). 2016. Living planet report. WWF International.
Young, A., Boyle, T. & Brown, T. 1996. The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution,* 11, 413-418.

Steane, D. A., Potts, B. M., Mclean, E. H., Collins, L., Holland, B. R., Prober, S. M., Stock, W. D., Vaillancourt, R. E. & Byrne, M. 2017. Genomic scans across three eucalypts suggest that adaptation to aridity is a genome-wide phenomenon. *Genome biology and evolution,* 9**,** 253-265.

Stingemore, J. A. & Krauss, S. L. 2013. Genetic delineation of local provenance in Persoonia longifolia: implications for seed sourcing for ecological restoration. *Restoration Ecology,* 21**,** 49-57.

Taylor, R.S., Watson, S.J., Bennett, A.F. and Clarke, M.F., 2013. Which fire management strategies benefit biodiversity? A landscape-perspective case study using birds in mallee ecosystems of south-eastern Australia. *Biological Conservation*, 159, pp.248-256.

Turesson, G. 1922. The genotypical response of the plant species to the habitat. *Hereditas,* 3**,** 211-350.

Verdone, M. & Seidl, A. 2017. Time, space, place, and the Bonn Challenge global forest restoration target. *Restoration Ecology*.

Violle, C., Navas, M. L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. 2007. Let the concept of trait be functional! *Oikos,* 116**,** 882-892.

Vranckx, G. U. Y., Jacquemyn, H., Muys, B. & Honnay, O. 2011. Meta-analysis of susceptibility of woody plants to loss of genetic diversity through habitat fragmentation. *Conservation Biology,* 26**,** 228-237.

Weiher, E. 2007. On the status of restoration science: obstacles and opportunities. *Restoration Ecology,* 15**,** 340-343.

Weiher, E., Werf, A., Thompson, K., Roderick, M., Garnier, E. & Eriksson, O. 1999. Challenging Theophrastus: a common core list of plant traits for functional ecology. *Journal of vegetation science,* 10**,** 609-620.

Wilkinson, D. M. 2001. Is local provenance important in habitat creation? *Journal of Applied Ecology,* 38**,** 1371-1373.

Williams, A. V., Nevill, P. G. & Krauss, S. L. 2014. Next generation restoration genetics: applications and opportunities. *Trends in Plant Science,* 19**,** 529-537.

Whitham, T.G., Bailey, J.K., Schweitzer, J.A., Shuster, S.M., Bangert, R.K., LeRoy, C.J., Lonsdorf, E.V., Allan, G.J., DiFazio, S.P., Potts, B.M. and Fischer, D.G., 2006. A framework for community and ecosystem genetics: from genes to ecosystems. *Nature Reviews Genetics*, *7*(7), pp.510-523.

World Wide Fund for Nature (WWF). 2016. Living planet report. WWF International.

Young, A., Boyle, T. & Brown, T. 1996. The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution,* 11**,** 413-418.

Statement of Authorship

Title of Paper	Increased genetic diversity via gene flow provides hope for Acacia whibleyana, an endangered wattle facing extinction
Publication Status	■ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Blyth, C., Christmas, M.J., Bickerton, D.C., Faast, R., Packer, J.G., Lowe, A.J. and Breed, M.F., 2020. Increased genetic diversity via gene flow provides hope for Acacia whibleyana, an endangered wattle facing extinction. *Diversity*, *12*(8), p.299.

Principal Author

Name of Principal Author (Candidate)	Colette Blyth		
Contribution to the Paper	Helped develop the ideas and methods, conducted fieldwork, analysed the data and wrote the manuscript.		
Overall percentage (%)	60		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	8/7/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr Matt Christmas		
Contribution to the Paper	Helped develop the ideas and methods, helped analyse the data and contributed to draft revisions of manuscript.		
Signature		Date	8/7/2021
Name of Co-Author	Doug Bickerton		

Contribution to the Paper	Helped develop the ideas and methods, helped with fieldwork, contributed to draft revisions of manuscript.		
Signature		Date	2/7/21
Name of Co-Author	Dr Renate Faast		
Contribution to the Paper	Helped develop the ideas and methods, helped with fieldwork, contributed to draft revisions of manuscript.		
Signature		Date	4th July 2021
Name of Co-Author	Dr Jasmin Packer		
Contribution to the Paper	Helped develop the ideas and methods, helped with fieldwork, contributed to draft revisions of manuscript.		
Signature		Date	8/7/2021
Name of Co-Author	Professor Andy Lowe		
Contribution to the Paper	Helped secure funding, helped develop the ideas and methods, contributed to draft revisions of manuscript.		
Signature		Date	5/7/2021
Name of Co-Author	Dr Martin Breed		
Contribution to the Paper	Helped secure funding, helped develop the ideas and methods, helped analyse the data, contributed to draft revisions of manuscript.		
Signature		Date	2/7/2021

Chapter 2. Increased Genetic Diversity via Gene Flow Provides Hope for *Acacia whibleyana*, an Endangered Wattle Facing Extinction

diversity

MDPI

Article

Increased Genetic Diversity via Gene Flow Provides Hope for *Acacia whibleyana*, an Endangered Wattle Facing Extinction

Colette Blyth [1,2,*], Matthew J. Christmas [3,*], Doug C. Bickerton [4], Renate Faast [1], Jasmin G. Packer [1,5], Andrew J. Lowe [1] and Martin F. Breed [2,6]

1 School of Biological Sciences, University of Adelaide, North Terrace, SA 5005, Australia; renate.faast@adelaide.edu.au (R.F.); j.packer@adelaide.edu.au (J.G.P.); andrew.lowe@adelaide.edu.au (A.J.L.)
2 Group on Earth Observations Biodiversity Observation Network (GEO BON) Genetic Composition Working Group, 04103 Leipzig, Germany; martin.breed@flinders.edu.au
3 Department of Medical Biochemistry and Microbiology, Uppsala University, Uppsala 753 41, Sweden
4 Northern Territory Herbarium, Palmerston, NT 0831, Australia; dugalugs29@gmail.com
5 Environment Institute, University of Adelaide, North Terrace, SA 5005, Australia
6 College of Science and Engineering, Flinders University, Bedford Park, SA 5042, Australia
* Correspondence: colette.blyth@adelaide.edu.au (C.B.); matthew.christmas@imbim.uu.se (M.J.C.)

Received: 29 June 2020; Accepted: 24 July 2020; Published: 30 July 2020

check for updates

Abstract: In this paper we apply a conservation genomics approach to make evidence-based management recommendations for *Acacia whibleyana*, an endangered shrub endemic to Eyre Peninsula, South Australia. We used population genomic analysis to assess genetic connectivity, diversity, and historical inbreeding across all known stands of the species sampling remnant stands, revegetated stands of unknown origin, and a post-fire seedling cohort. Our results indicate a degree of historical connectivity across the landscape, but habitat loss and/or pollinator community disruption are potential causes of strong genetic structure across the remnant stands. Remnant stands had low genetic diversity and showed evidence of historical inbreeding, but only low levels of intra-stand relatedness indicating that risks of contemporary inbreeding are low. Analysis of a post-fire first generation cohort of seedlings showed they likely resulted from intra-stand matings, resulting in reduced genetic diversity compared to the parents. However, admixed seedlings in this cohort showed an increase in heterozygosity relative to likely sources and the non-admixed seedlings of the same stand. Assisted inter-stand gene flow may prove an effective management strategy to boost heterozygosity and corresponding increases in adapting capacity in this endangered species.

Keywords: conservation genetics; endangered species; genetic diversity

1. Introduction

The likelihood that a species is able to adapt to both short-term environmental variability and long-term change is positively correlated with its level of genetic diversity [1]. Small and/or fragmented populations often suffer from restricted gene flow and small effective population sizes, which increase the effects of genetic drift. In small populations, the chances of inbreeding are high, resulting in further reductions in genetic diversity. Genetic drift and inbreeding are therefore potential drivers of the extinction vortex and reduce the adaptive potential of small populations [1–6]. The maintenance and, where possible, enhancement of genetic diversity is therefore a critical aim of conservation genetic strategies and can be achieved through increasing connectivity and gene flow between subpopulations [7].

Connecting isolated populations via the translocation of individuals involves an element of risk [7]. Admixture may in fact lead to reduced population fitness due to outbreeding depression, where the introduction of non-local genotypes can disrupt local adaptation and/or lead to the breakdown of adaptive gene complexes [8]. In the case of rare and threatened species, however, the risks of outbreeding depression are often outweighed by the consequences of inbreeding and low adaptive potential if no action is taken [7,9,10]. Indeed, the primary risks for threatened species, especially those that occur in small and fragmented populations, are the deleterious inbreeding and genetic drift effects [9,10]. In such situations, the introduction of non-local genotypes may instead increase a population's chances of survival due to a genetic rescue effect [11–13].

The development of high throughput sequencing technologies over the last decade has facilitated research into conservation genetics in plant species, with an aim to inform conservation practice on the appropriate management of genomic diversity for improved conservation outcomes [14–16]. Such studies can provide insights into the genetic diversity within plant populations, and the historical gene flow between populations, as well as identifying where genetic breaks may occur [16]. This information can be used to guide when and how to enhance the genetic diversity of an endangered species while minimising the risks associated with moving genetic material across large genetic breaks in a species range, such as outbreeding depression [16]. As a result of such studies, there have been increasing calls for a paradigm shift in conservation management thinking—away from the separate management of isolated populations and towards a strategy that seeks to increase connectivity and diversity across the range of a species in the hope it will lead to genetic rescue [13,17].

First described in the 1990s, the endangered *Acacia whibleyana* (family Fabaceae) is restricted to the Tumby Bay area of the Eyre Peninsula, South Australia. There have been significant investments to aid its conservation, and it is part of the Australian Government's *30 by 2020 Threatened Species Strategy* [18]. Land clearing has limited its suitable habitat, and ongoing low seed set and lack of natural recruitment, as well as competition from weeds, are some of its key threatening processes [19–21]. One of the priority conservation actions for *A. whibleyana* is to establish new translocated populations [19–21]. Conservation management decisions should ideally be underpinned by a sound understanding of its genetic structure. However, to date, there has been no published research on the levels of genetic variation and structure within this species.

Here, we address this knowledge gap by using high throughput sequencing to construct a high-resolution picture of the population genetic structure and diversity in *A. whibleyana*. We generated a genome-wide single nucleotide polymorphism (SNP) dataset for the last remaining stands of *A. whibleyana,* including all remnant and revegetated stands, plus a cohort of seedlings that recruited post-fire. We used this population genomic information to provide practical conservation management recommendations for the conservation of this species.

2. Materials and Methods

2.1. Study Species and Sampling

Acacia whibleyana is a narrow-ranged endemic shrub that is restricted to five subpopulations near Tumby Bay, Eyre Peninsula, South Australia (Figure 1). The species is not known to hybridise [21].

Figure 1. Map of collection stands. The Southerns and Marshalls Rd locations had collections of more than one status (i.e., a combination of remnant, revegetated and/or seedling plants). The figure was generated using Quantum Geographic Information System (Quantum GIS Development Team, 2016, Open source geospatial foundation project, http://qgis.osgeo.org). Map layer was downloaded from https://www.diva-gis.org/Data.

Acacias have an important nitrogen fixation and nutrient cycling role in many ecosystems, which is of particular ecological importance in the predominantly arid, nutrient poor soils of Australia [22]. It produces flowers between August and October [23], with fruits maturing from December to January. It has four to six seeds per pod, and seed dispersal is believed to be assisted by ants [23]. Like other wattle species, the seeds of *A. whibleyana* have a hard seed-coat dormancy mechanism that requires disturbance to scarify the seed and trigger germination.

A species census in April 2019 recorded ~1800 individuals across all known subpopulations. The maximum geographic distance between subpopulations was 38 km. The total area of occupancy of *Acacia whibleyana* was 14.2 ha (0.142 km^2) based on site-validated data of each remnant stand. Many plants were restricted to remnant vegetation on roadside verges, with some found on privately-owned land which is grazed by livestock [21] (Table A1 in Appendix A). One of the existing stands (Marshalls Rd) was established in 2004 through two separate translocations, with plants from the Quarry stand. The oldest Marshalls Rd stand was experimentally burned in October 2018 to explore the effect of fire on recruitment.

Initially, we collected *A. whibleyana* samples in late 2017. Sampling was rangewide and included 5 remnant stands and a revegetated stand. On average, we aimed to collect 20 individuals per stand where possible to allow for extraction and genotyping failures, but this varied depending on census size and accessibility (Table A1). We targeted young leaves to maximise DNA quality, and once picked, leaves were placed in gauze bags and desiccated on silica gel until DNA extraction.

In May 2018, a prescribed burn of the revegetated Marshalls Rd stand was done to encourage natural regeneration. An additional 46 seedlings were collected in October 2018 post-burn for genotyping. Seedling leaves were placed directly into DNA extraction plates for processing of the fresh leaf material.

In February 2019, an additional sampling trip took place, at which time samples were collected from both a newly-discovered stand of *A. whibleyana* within the Salt Lake subpopulation and from a known stand previously not sampled. Samples were also collected from a site on the western edge of the Quarry subpopulation referred to as Southerns.

2.2. DNA Extraction, Sequencing and Filtering

DNA extraction, library preparation, and sequencing were done at the Australian Genome Research Facility. We used the Machery-Nagel Nucleospin Plant II Kit for DNA extraction. Double digest restriction-associated DNA sequencing (ddRADseq; [24] was used to generate reduced representations of the genome for SNP calling across all samples. Full details of the method can be found in Peterson, Weber [24]. Briefly, ddRAD library preparation protocol involved the following steps: (1) DNA digestion with two restriction enzymes (PstI and MseI); (2) ligation of barcoded adaptors, specific to each sample, to restriction site overhangs; (3) a wide size selection of pooled digested-ligated fragments 280–342bp using Blue Pippin (Sage Science, Beverly, MA, USA) (Wide); (4) amplification of library via polymerase chain reaction (PCR) for 11 cycles using indexed primers. The libraries were assessed by gel electrophoresis (Agilent D1000 ScreenTape Assay), quantified by quantitative polymerase chain reaction (qPCR; KAPA Library Quantification Kits for Illumina) and then sequenced with 150 bp single reads on the NextSeq 500 system using NextSeq 500 high Output Kit v2 (150 cycles) reagents.

Following sequencing, reads were processed with the STACKS pipeline [25,26] at Australian Genome Research Facility(AGRF; Melbourne, Australia). Briefly, reads were deconvoluted by inline barcodes, checked for read quality, and for the presence of a restriction site, creating per-sample FASTQ files. These FASTQ files were then trimmed to the size of the shortest read minus 2 bp to compensate for differences in read length due to any variation in barcode length. After trimming, stacks of similar reads were created for each sample individually; these read stacks are also known as tags. Tags which appear across all samples were collated (catalogue tags), and genotypes were then allocated to the common polymorphic sites. The collated SNPs across all individuals were then filtered with the following settings: minimum number of reads required at a stack to call a homozygous genotype = 5; minor allele frequency, below which a stack is called a homozygote = 0.05; minimum minor allele frequency to call a heterozygote = 0.1 (above 0.05 but below 0.1, a stack is called 'unknown'). Computed genotypes were then exported as a single .vcf file for further filtering and analysis.

We used VCFtools [27] to remove SNPs that had a minimum depth of coverage of <10 reads per individual or a minor allele frequency of <0.1. VCFtools was also used to assess missing data, removing SNPs if they had >25% missing data across all individuals and removing individuals if they had >25% missing SNP calls. The –indep-pairwise command in PLINK [28] was then used to identify SNPs in linkage disequilibrium, whereby one SNP in a pair of linked SNPs ($r^2 > 0.8$) was pruned from the data set.

Inbreeding coefficient (F) values were estimated using the G-stats calculator in GENODIVE v. 3.2 [29] for each SNP across all samples, and SNPs with significantly negative F removed using VCFtools (significance assessed using permutation tests with 10,000 permutations and a significance threshold of $p < 0.05$). Negative F indicates greater than expected heterozygosity under Hardy-Weinberg equilibrium, which may be indicative of paralogous reads being stacked together in the STACKS pipeline [30].

2.3. SNP Calling

We generated 467,346,192 sequences across the 237 *A. whibleyana* samples in our study. A total of 281,779 SNPs were called after read alignment. After filtering on depth of coverage, minimum minor allele frequency, percentage of missing data, linkage disequilibrium, 16,776 SNPs across 228 individuals remained in our dataset (see Table 3 for final sample sizes). Computed genotypes were then exported as a single .vcf file for further filtering and analysis (Supplementary file S1). The final, filtered SNP set was then used to investigate population genetic structure, diversity, and inbreeding.

2.4. Genetic Structure Analysis

We estimated pairwise genetic differentiation (F_{ST}) between all 11 stands (including the remnant, revegetated, translocated, and the seedling cohort) in GENODIVE v3.2 [29]. We also used two different genetic clustering analyses to identify the most likely number of genetic clusters (K) in our dataset. Firstly, we used the non-model based Discriminant Analysis of Principal Components (DAPC) using the R package adegenet [31]. DAPC is a multivariate approach that is not based on population genetic models, but rather seeks discriminating functions between groups of individuals while minimising variation within clusters. Genetic data were first transformed into uncorrelated components using principal component analysis (PCA). The number of genetic clusters was then defined using k-means, a clustering algorithm that looks for the value of K that maximises the variation between groups. The Bayesian Information Criterion (BIC) was calculated for K = 1 to 20, and the K value with the lowest BIC was selected as the optimal number of clusters. A discriminant analysis was then performed on the first 80 principal components using the function dapc, implemented in R, in order to efficiently describe the genetic clusters and assign samples to each cluster.

The model-based genetic clustering algorithm ADMIXTURE [32] was also used to estimate the most likely number of clusters (K) in our dataset. ADMIXTURE includes a cross-validation procedure for identifying the value of K for which the model has best predictive accuracy (see ADMIXTURE manual for full details of model and cross validation procedure). ADMIXTURE was also run for K values 1 to 10 and the most likely value of K was assessed by comparing the cross-validation errors (cv errors) between runs, with the lowest cv error indicating greatest support. We repeated each ADMIXTURE analysis ten times and used CLUMPAK to combine the results and construct bar plots of individual assignment to clusters for the most supported values of K using DISTRUCT [33,34].

GENODIVE v3.2 [29] was used to run a nested analysis of molecular variance (AMOVA) to explore genetic variation within individuals, among individuals nested within collection stand, and among collection stands.

2.5. Genetic Diversity and Inbreeding Analysis

We used GENODIVE to estimate the inbreeding coefficient (F), and observed (H_O) and expected (H_E) heterozygosity. We also ran these analyses on only Marshalls Rd and the Marshalls Rd seedlings, comparing admixed and non-admixed individuals (as identified from the DAPC and ADMIXTURE analyses) in the seedling cohort. In order to compare levels of heterozygosity in the Marshalls Rd seedlings that were identified as admixed and non-admixed (see results), we calculated individual heterozygosity using the -het flag in VCFTools. These values were then imported into R for plotting using the ggplot2 package [35]. To further explore historical inbreeding and future risks of inbreeding, we estimated pairwise kinship coefficients in GENODIVE. Kinship varies between 0 and 1, where 0 indicates no genetic evidence of recent co-ancestry and 1 indicates samples are genetically identical (i.e., clones).

2.6. Redundancy Analysis

All stands were located in close proximity to each other (the maximum distance between any two stands was 38 km). We ran a redundancy analysis (RDA) on the allele frequencies of all SNPs in order to assess how much of the genetic variation among our populations can be explained by spatial factors (i.e., resulting from isolation by distance) using the VEGAN analytical package v2.5-6 in R (https://cran.r-project.org/web/packages/vegan/). First, spatial coordinates for each sampling stand were centred and converted into third degree polynomials. An RDA of allele frequencies~all spatial polynomials was then performed using the 'rda' function. Analysis of variance (ANOVA) was used to assess the significance of the RDA using the 'anova.cca' function.

3. Results

3.1. Genetic Structure Analysis

Both the DAPC and ADMIXTURE analyses indicated significant genetic structure in our data (Figure 2). The BIC values from DAPC revealed similar levels of support for K = 3, 4, or 5. Cross validation error values from the ADMIXTURE analysis showed greatest support for K-values of 5–7. However, beyond K = 5 the ADMIXTURE results appeared noisy and clustering closely matched individual source population for K values below six.

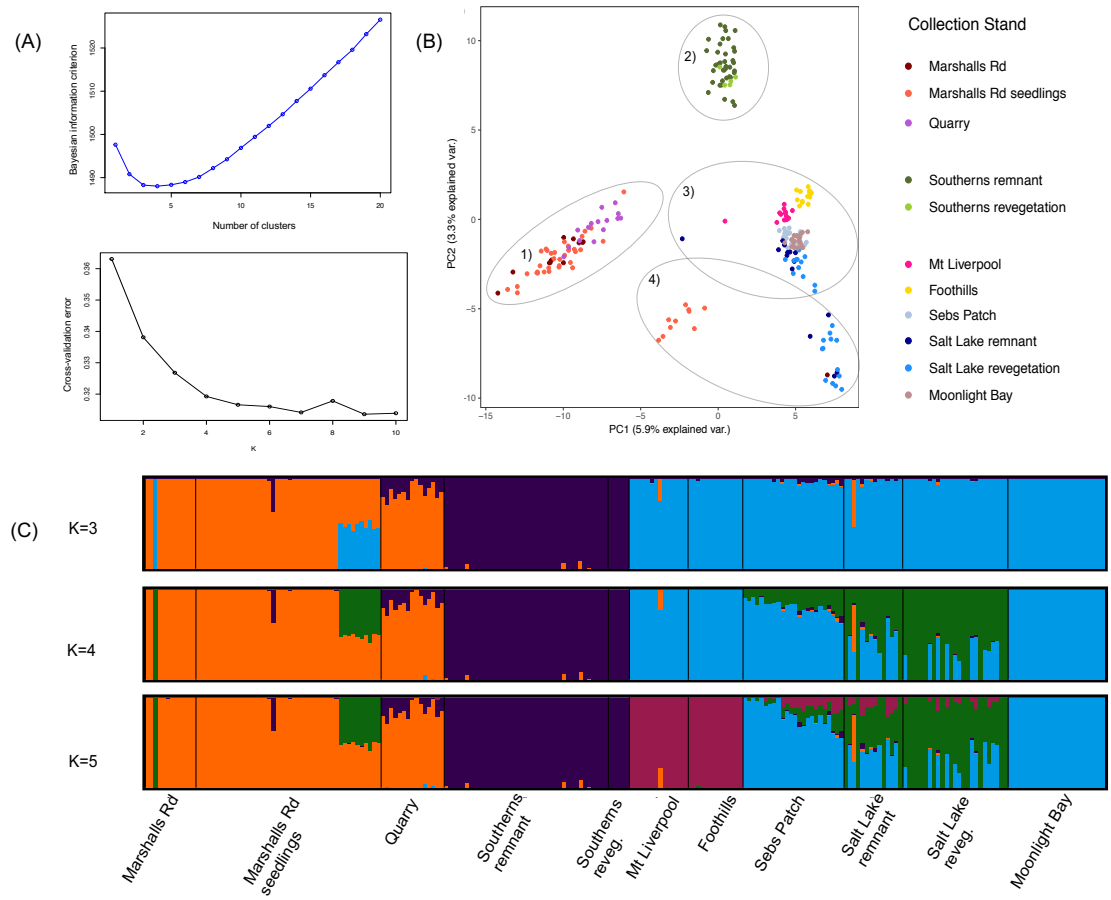


Figure 2. (**A**) The Bayesian Information Criterion (BIC) values for each number of clusters, with the lowest BIC indicating support for 3 or 4 clusters plus the ADMIXTURE cross-validation error (cv error) results. The cv errors are plotted for K = 1–10, with the lowest cv error indicating greatest support at K = 7. (**B**) Principal Component Analysis showing population genetic structure across all samples. Colours represent stands and solid line ellipses identify the genetic clusters (1–4) identified in the Discriminant Analysis of Principal Components (DAPC) analysis for K = 4. Circle with the dashed line highlights the seedlings that grouped with a different genetic cluster (cluster 4) to the Marshalls Rd mature plans and non-admixed seedlings (cluster 1). (**C**) Bar plots represent individual genetic cluster assignment from ADMIXTURE results from K = 3 to K = 5 (* note-beyond K = 5 the ADMIXTURE results appeared noisy but see Figure A1 for full output).

At K = 3, individuals were clustered into groups that aligned with geography: a northern cluster containing samples from Quarry and all but the one above-mentioned Marshalls Rd individual (Cluster 1), and a second cluster containing all samples from Southerns (Cluster 2), and a southern cluster, containing all individuals collected from stands in the south (Salt Lake reveg. and remnant, Sebs Patch, Moonlight Bay, Foothills Rd, and Mt Liverpool, and one individual from Marshalls Rd; Cluster 3). In the ADMIXTURE analysis, ten of the Marshalls Rd seedlings showed approximately 50/50 admixture between Clusters 1 and 3 and, in the DAPC analysis, clustered with Cluster 3.

For K = 4, substructure was revealed within the revegetated and remnant Salt Lake stands, with 17 individuals having almost 100% assignment to this 4th cluster (Cluster 4), with the rest displaying admixture between clusters 1 and 4. Individuals from the nearby Sebs Patch stand displayed a low level of admixture with Cluster 4 (average assignment 16%). The Marshalls Rd individual that showed 100% assignment to Cluster 3 when K = 3 had 100% assignment to Cluster 4, and the admixed seedlings showed 50% assignment to Cluster 4. Further substructure was revealed for K = 5, in Foothills Rd and Mt Liverpool, with the Foothills Rd and Mt Liverpool stands forming a distinct genetic cluster (Cluster 5) which distinguished them from Moonlight Bay plants.

The majority of the Marshalls Rd seedlings clustered with the Quarry and Marshalls Rd samples in all genetic structure analyses. However, of the 46 seedlings genotyped, 10 of the samples showed ca. 50/50 ancestry with Marshalls Rd and Salt Lake in the ADMIXTURE analysis, suggesting that they are likely to be offspring from a cross between a Marshalls Rd individual and a Salt Lake individual. These seedlings were also located halfway between the Marshalls Rd samples and Salt Lake samples in the PCA, and clustered with some Salt Lake revegetation and remnant samples in the DAPC (Figure 1).

The genetic proximity of clusters generally reflected geographic proximity. This was supported by results of pairwise genetic differentiation (F_{ST}) between stands (Table 1) and the RDA. Pairwise F_{ST} between stands was reasonably high (a maximum of 0.193, Table 1). However, there was high differentiation between the two geographically close stands Foothills and Mt Liverpool (pairwise F_{ST} = 0.185, Table 1). These stands also displayed comparatively high genetic differentiation from all other stands (Mt Liverpool pairwise F_{ST} = 0.100 to 0.185 and Foothills pairwise F_{ST} = 0.129 to 0.193, Table 1). The redundancy analyses of allele frequencies ~ space revealed that the spatial variables explained 83% of the variation in allele frequencies among stands (ANOVA, F = 6.43, p = 0.017), showing spatial variation to be a strong explanatory variable of the observed patterns of genetic differentiation.

Results showed all historical revegetated stands clustered with their nearby remnant stands. Marshalls Rd (revegetated) and Quarry (remnant) were genetically very similar (Pairwise F_{ST} = 0.017, Table 1) and clustered together in PCA, DAPC, and ADMIXTURE analyses, most likely reflecting the use of the Quarry stand as the source population for the Marshalls Rd stand. The samples taken from Salt Lake revegetation were genetically similar to samples taken from the remnant individuals at Salt Lake (Pairwise F_{ST} = 0.000; Table 1.) and also clustered together in the PCA, DAPC, and ADMIXTURE analyses. Similarly, the samples collected from the Southerns revegetation stand were genetically similar to those from the Southerns remnant stand (Pairwise F_{ST} = 0.013; Table 1.), again corroborated by the PCA, DAPC and ADMIXTURE analyses.

Analysis of molecular variance revealed that the majority of genetic variation is found among all individuals (73.6%). Significant levels of genetic variation are also found among individuals within a collection stand (15.7%, $p < 0.001$), as well as between collection stands (10.7%, $p < 0.001$, Table 2), as highlighted by the genetic structure analysis.

Table 1. Pairwise genetic differentiation (measured by F_{ST}) between stands (values vary from 0-1 and values of 0 indicate no evidence of differentiation).

Collection Stand	*Marshalls Rd*	*Marshalls Rd Seedlings*	*Quarry*	*Southerns Remnant*	*Southerns reveg.*	*Mt Liverpool*	*Foothills*	*Sebs Patch*	*Salt Lake remnant*	*Salt Lake reveg.*	*Moonlight Bay*
Marshalls Rd	0										
Marshalls Rd seedlings	−0.008	0									
Quarry	0.017	0.032	0								
Southerns remnant	0.090	0.101	0.087	0							
Southerns reveg.	0.101	0.114	0.101	0.013	0						
Mt Liverpool	0.156	0.158	0.153	0.122	0.145	0					
Foothills	0.192	0.193	0.189	0.147	0.176	0.185	0				
Sebs Patch	0.109	0.117	0.111	0.083	0.090	0.104	0.134	0			
Salt Lake remnant	0.088	0.099	0.097	0.073	0.077	0.100	0.129	0.037	0		
Salt Lake reveg.	0.115	0.123	0.126	0.097	0.104	0.122	0.148	0.061	0.000	0	
Moonlight Bay	0.136	0.145	0.137	0.111	0.116	0.135	0.161	0.084	0.075	0.097	0

Table 2. Nested analysis of molecular variance (AMOVA) including standard deviations (SD; obtained through jackknifing over loci) and 95% confidence intervals (c.i. 95%; obtained through bootstrapping over loci) of *F* statistics.

Source of Variation	*Nested in*	*%var*	*F-Value*	*SD*	*c.i.95%*
Within Individual	-	73.6	0.263	0.002	±0.004
Among Individual	Collection stand	15.7	0.175	0.002	±0.004
Among Population		10.7	0.107	0.001	±0.002

3.2. Genetic Diversity and Inbreeding

Expected heterozygosity varied significantly but by a small degree between stands (H_E = 0.117 to 0.141, Table 3). Observed heterozygosity was consistently lower and less variable than expected heterozygosity (H_O = 0.102 to 0.118, Table 3), which is reflected in the positive inbreeding coefficients (F) found in all collection stands (F ranged from 0.110 to 0.226; Table 3), indicating the likely presence of historical inbreeding. Mt Liverpool and Foothills had significantly lower inbreeding coefficients than other stands (F = 0.11, 0.13 respectively, Table 3). Moonlight Bay, Sebs Patch, and Salt Lake remnant stands had the highest inbreeding coefficients (F = 0.226 to 0.211, Table 3). The confidence intervals from the kinship analysis spanned zero for every stand, suggesting that relatedness, and therefore contemporary inbreeding, was low within stands (Table 3).

We found significant differences in H_E, H_O and F between the mature Marshalls Rd plants and both the non-admixed and admixed seedlings (Table 3, Figure A2). The non-admixed seedlings displayed lower heterozygosity and higher F compared to the mature plants (Table 3, Figure A2). In contrast, the admixed seedlings had higher heterozygosity compared to both the non-admixed seedlings and mature plants (Table 3, Figure A2). Further, we found the inbreeding coefficient was significantly lower in the admixed seedlings compared to the non-admixed seedlings (F = 0.036 and 0.181 respectively, Table 3, Figure A2).

Table 3. Stand status, sample size after filtering (n), census population size, location, observed (H_O) and expected (H_E) heterozygosity (both vary from 0 to 1, with 0 indicating no diversity), inbreeding coefficient (F) (varies from 0 to 1 and values of 0 indicate no evidence of inbreeding), and kinship (varies between 0 and 1, where 0 indicates no recent co-ancestry and 1 indicates samples are genetically identical). Includes a focused analysis for both the admixed and non-admixed Marshalls Rd seedlings.

Collection Stand	*Status*	*n*	H_O	H_E	*F*	*Kinship*
All stands						
Marshalls Rd	Translocation	12	0.114 (0.111 to 0.116) ^{b,c}	0.137 (0.135 to 0.140) ^{c,d}	0.173 (0.163 to 0.182) ^{b}	−0.008 (0.018 to −0.027) ^{a,b}
Marshalls Rd seedlings (all)	Seedlings	44	0.107 (0.105 to 0.109) ^{a,b}	0.131 (0.128 to 0.133) ^{b}	0.183 (0.176 to 0.189) ^{b}	0.001 (0.003 to −0.002) ^{a}
Quarry	Remnant	15	0.111 (0.109 to 0.114) ^{b}	0.136 (0.134 to 0.139) ^{c,d}	0.183 (0.174 to 0.192) ^{b}	0.004 (0.011 to −0.003) ^{a}
Southerns remnant	Remnant	39	0.118 (0.116 to 0.121) ^{c}	0.145 (0.142 to 0.147) ^{d}	0.183 (0.177 to 0.189) ^{b}	0.000 (0.003 to −0.003) ^{a}
Southerns reveg.	Revegetation	5	0.118 (0.115 to 0.121) ^{c}	0.140 (0.137 to 0.143) ^{c,d}	0.157 (0.143 to 0.170) ^{b}	0.006 (0.043 to −0.031) ^{a,b}
Mt Liverpool	Remnant	14	0.109 (0.107 to 0.112) ^{b}	0.123 (0.120 to 0.126) ^{a}	0.110 (0.100 to 0.121) ^{a}	0.009 (0.026 to −0.007) ^{a}
Foothills	Remnant	13	0.102 (0.099 to 0.105) ^{a}	0.117 (0.114 to 0.120) ^{a}	0.130 (0.119 to 0.141) ^{a}	0.008 (0.023 to −0.006) ^{a}
Sebs Patch	Remnant	24	0.109 (0.107 to 0.111) ^{b}	0.138 (0.136 to 0.141) ^{c,d}	0.211 (0.204 to 0.218) ^{c}	0.001 (0.006 to −0.003) ^{a}
Salt Lake remnant	Remnant	14	0.110 (0.108 to 0.112) ^{b}	0.141 (0.139 to 0.144) ^{d}	0.221 (0.212 to 0.229) ^{c}	−0.011 (−0.008 to −0.014) ^{b}
Salt Lake reveg.	Revegetation	25	0.117 (0.115 to 0.120) ^{b,c}	0.138 (0.136 to 0.141) ^{c,d}	0.151 (0.144 to 0.159) ^{b}	0.001 (0.009 to −0.006) ^{a}
Moonlight Bay	Remnant	23	0.104 (0.102 to 0.106) ^{a}	0.135 (0.132 to 0.137) ^{b,c}	0.226 (0.219 to 0.234) ^{c}	−0.004 (0.000 to −0.008) ^{a,b}
Marshalls Rd focus						
Marshalls Rd adults	Translocation	12	0.114 (0.111 to 0.116) ^{b}	0.137 (0.135 to 0.140) ^{b}	0.173 (0.163 to 0.182) ^{b}	−0.008 (0.018 to −0.027) ^{a}
Marshalls Rd seedlings (non-admixed only)	Seedlings	34	0.103 (0.101 to 0.106) ^{a}	0.126 (0.124 to 0.129) ^{a}	0.181 (0.173 to 0.188) ^{b}	0.008 (0.010 to −0.002) ^{a}
Marshalls Rd seedlings (admixed only)	Seedlings	10	0.120 (0.117 to 0.123) ^{c}	0.125 (0.122 to 0.127) ^{a}	0.036 (0.023 to 0.049) ^{a}	0.117 (0.122 to −0.005) ^{a}

95% confidence intervals in parentheses and homogeneous subgroups indicated by ^{a}, ^{b}, ^{c}, and $^{d.}$

4. Discussion

Our analysis of genomic diversity and differentiation among the remaining stands of the endangered *Acacia whibleyana* has revealed many of the hallmarks of a species trapped in an extinction vortex. There was strong genetic structure among stands, indicating that gene flow between stands is likely restricted and genetic drift has driven divergence among stands. Positive inbreeding coefficients across all stands suggest the presence of historical inbreeding, however analysis of kinship shows the relatedness of remaining plants to be low within stands, suggesting the risks of contemporary inbreeding are low and potentially that this species has mechanisms for inbreeding avoidance. We found that seedlings resulting from admixture events between the distinct genetic clusters had higher genetic diversity compared to progeny of single-parent stands. Together, these results indicate that while the species presents low risks of contemporary inbreeding, the establishment of mixed genetic cluster stands may be required to boost the genetic diversity of subsequent generations and provide an escape route from the extinction vortex this species is facing.

4.1. Genetic Structure and Sources of Genetic Variation

The level of genetic structure found among *A. whibleyena* stands was high, given the small geographic range of the species (<38 km), and spatial location was a strong explanatory variable of the genetic variation in the species. The fact that the remaining *A. whibleyana* plants were distributed across small subpopulations suggests that genetic drift has played a part in shaping the genetic differentiation of the stands, as has been found for other endangered plant species, including the comparable *Acacia pinguifolia* [36]. Acacias are mostly pollinated by generalists, with bees their most common pollinators [37]. Concerns have been raised over the generalised threat of pollinator limitation to genetic diversity and inbreeding [38], with supporting evidence from other endangered Acacias [37]. Pollinator limitation could, therefore, be part of the explanation for the high genetic structure and low levels of genetic diversity found in *A. whibleyana* and may represent a major threat to the long-term survival of the species. Another possible contributing factor is that the seed is predominantly dispersed by ants and therefore only over short distances [39]. This, coupled with the lack of alternative suitable habitats for seed to disperse to, may also limit gene flow, however pollen dispersal is thought to have a far greater effect on genetic structure [40].

Interestingly, we found strong genetic structure among stands that were closely located. The Southerns site was found to be significantly differentiated from the Marshalls Rd and Quarry sites despite its proximity to them (<1 km). Further, the highest level of genetic differentiation was between two of the most geographically proximate stands—Foothills and Mt Liverpool. Despite the high F_{ST} between these stands reflecting large differences in allele frequencies, they cluster together and separately from all other sites at K = 5 in the ADMIXTURE analysis, suggesting that they share variation not found elsewhere. The small population size of these stands is likely to have led to a pronounced effect of genetic drift increasing relative divergence between them.

4.2. Opportunistic Test of Admixture on Genetic Diversity

We present evidence that seedlings at Marshalls Rd likely resulted from crosses between individuals from distinct genetic clusters. The likely source of this admixture is a single Marshalls Rd adult that clusters with Salt Lake individuals. This individual is likely to have originated from a translocation event, such as from seed from the Salt Lake site that was included in the original Quarry/Marshalls Rd translocation project. Plants were propagated for revegetation projects at both sites in the same nursery during this time, so we believe this genetic transfer is likely (Geraldine Turner, pers. Comm, May 2020.). While we cannot rule out the presence of other Salt Lake-like individuals at the Marshalls Rd site, as we did not genotype all mature plants and seedlings, the presence of at least one Salt Lake-like individual is strongly indicative of an occurrence of within-stand mating resulting in the admixed seedlings, rather than gene flow between the geographically distant Marshalls Rd and Salt Lake stands.

The admixed seedlings provided us with an unexpected opportunity to explore the effects of inter-population crossing on heterozygosity and inbreeding. When comparing the admixed to the non-admixed seedlings, our results show a striking reduction in the inbreeding coefficient and increase in observed heterozygosity in the admixed seedlings (admixed vs. non-admixed seedlings: F = 0.181 vs. 0.036; Ho = 0.103 vs. 0.120; He = 0.126 vs. 0.125). These large differences suggest that including individuals from a non-local population could bolster the genetic diversity of a population by encouraging matings between genetically dissimilar individuals. We also saw reduced observed heterozygosity in the non-admixed seedlings compared to mature Marshalls Rd plants, indicating that the species is experiencing a reduction in genetic diversity across generations, likely due to increased inbreeding, pushing the species further down the extinction vortex. Alternatively, these results could indicate that the seedlings included in our study were more inbred than would normally make it to maturity, and that insufficient time has passed for inbreeding depression to manifest.

Unfortunately, all of the seedlings genotyped in this study died, and so the fitness effects of higher or lower genetic diversity in the admixed and non-admixed seedlings respectively is unknown. Whilst it is promising to see that the admixed progeny display increased levels of heterozygosity, whether assisted gene flow among genetic clusters would prove an effective conservation strategy for increasing fitness in future generations is unclear. We see this as the next research goal for this species, with controlled crosses and genotype-fitness associations required to test these ideas.

4.3. Genetic Diversity and Inbreeding

Levels of heterozygosity were similar across all stands and observed heterozygosity was consistently lower than expected heterozygosity in all stands, which is indicative of historical inbreeding. It is difficult to assess whether or not genetic diversity is low in this species in the absence of a reference species, ideally a more widespread *Acacia* species. To our knowledge, no other studies using genome-wide SNP markers have been carried out on *Acacia* species. However, studies on other predominantly outcrossing tree species that have utilised SNP markers report heterozygosity values of between 0.1 and 0.4 [41–45], suggesting that *A. whibleyana* is at the lower end of this range. Additionally, the increase in heterozygosity that we observe in the admixed seedlings demonstrate that the genetic diversity we observe is not evenly distributed among the stands and can potentially be boosted in future generations via admixture. However, the fitness consequences of this boost in genetic diversity requires further exploration.

Australian *Acacia* species are generally preferential outcrossers [46–48], and therefore are likely to display mechanisms of inbreeding avoidance. In a study of *Acacia myrtifolia* [46], it was found that as relatedness between individuals increased, the number of set seed pods decreased. Low seed set has been observed in some stands of *A. whibleyana* [49], however our kinship analysis revealed relatedness among individuals in stands to be low. We also show there to be substantial genetic variation within individuals, meaning that the risks of contemporary inbreeding are low. There has also been suggestion that some endangered *Acacia* are intrinsically rare or can at least adapt to having lowered genetic diversity and small population size by increasing the levels of selfing or, in extreme cases, reproduce vegetatively [50,51]. It is unclear if *A. whibleyana* could respond in this manner.

4.4. Management Recommendations

The results we present here demonstrate that the remaining stands of *A. whibleyana* exhibit low inter-stand gene flow with evidence that genetic drift, and potentially inbreeding, will draw the species further down the extinction vortex in subsequent generations. Our findings are supported by demographic observations of *A. whibleyana*, where low seed set, a lack of recruitment, and poor seedling survival have been identified as threats to the its survival [21]. Similar patterns were also observed for *A. pinguifolia,* which is another threatened species endemic to a similar region to *A. whibleyana. A. pinguifolia* presented high levels of historical inbreeding, strong fine-scale genetic structure, and limited

inter-stand gene flow [36]. Crossing between populations was recommended as a strategy to help reduce the risk of extinction in this species [36].

The boost to genetic diversity that we observed in the seedlings resulting from natural crossings between two of the genetic clusters provides hope for the species and a potential escape from the extinction vortex. These chance crossings demonstrate that genetic diversity can substantially increase within a single generation. However, since a significant proportion of the diversity found between the remaining stands is an ongoing concern, we suggest that a long-term genetic rescue trial should be established. This trial should focus on simultaneously determining the effects and trade-offs of inbreeding vs. outbreeding depression, heterosis as a result from inter-stand crosses, and pollination deficiency. This trial would require experimentally crossing a range of selfed, intra-stand, and inter-stand individuals, spanning the extent of genetic distance observed, and monitoring plants into adulthood. This genetic rescue trial could be efficiently combined with an assessment of pollination deficiency; this pollination/pollinator effect is important to determine as it may be contributing to the poor demographic rates observed [21,38].

5. Concluding Remarks

All too often, it is easy to lose optimism when managing an endangered species faced with extinction. Here, through an unexpected natural experiment, we have been given a glimmer of hope and guidance for the management of *Acacia whibleyana*. Through the analysis of the seedlings, we have evidence that mixing between genetically different stands can increase genetic diversity, at least for a period of time in the first generation. Further, our results suggest that without action, we may experience a decrease in genetic diversity with each generation. We hope that a cautious approach to genetic recue will help lessen the extinction risk faced by *Acacia whibleyana* by increasing its levels of genetic diversity and hopefully its adaptability in the face of environmental change.

Supplementary Materials: The following are available online at http://www.mdpi.com/1424-2818/12/8/299/s1, Supplementary file S1—The variant call format (vcf) file is available as supplementary (Acacia_whibleyana_paper.vcf). Reads and mapping files will be archived at the NCBI SRA (accession number TBA).

Author Contributions: Conceptualization, C.B., M.F.B., A.J.L. and D.C.B.; Data curation, C.B. and M.F.B.; Formal analysis, C.B., M.J.C. and M.F.B.; Funding acquisition, M.F.B., A.J.L. and C.B.; Methodology, C.B., M.F.B., M.J.C., A.J.L., D.C.B., J.G.P. and R.F.; Project administration, C.B., M.F.B., M.J.C., A.J.L, D.C.B., J.G.P. and R.F.; Resources, M.F.B.; Supervision, M.F.B., A.J.L. and M.J.C.; Visualization, C.B., M.F.B., and M.J.C.; Writing—Original draft, C.B.; Writing—Review & editing, C.B., M.F.B., M.J.C., A.J.L., D.C.B., J.G.P., R.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Commonwealth Government's Threatened Species Recovery Fund, Eyre Peninsula Natural Resources Management Board (EPNRMB), and the Australian Research Council (grant numbers DE150100542 and DP150103414). This project was supported by The Holsworth Wildlife Research Endowment & The Ecological Society of Australia.

Acknowledgments: The authors wish to thank Geraldine Turner and Seb Drewer from the Eyre Peninsula Natural Resources Management Board (EPNRMB) for providing their assistance and expertise throughout this study, and everyone else who assisted us with sample collection. Thanks also go to Sam Turner-Blyth for help early on in the writing process. Finally, a special thanks to Jeremy Austin for his assistance behind the scenes.

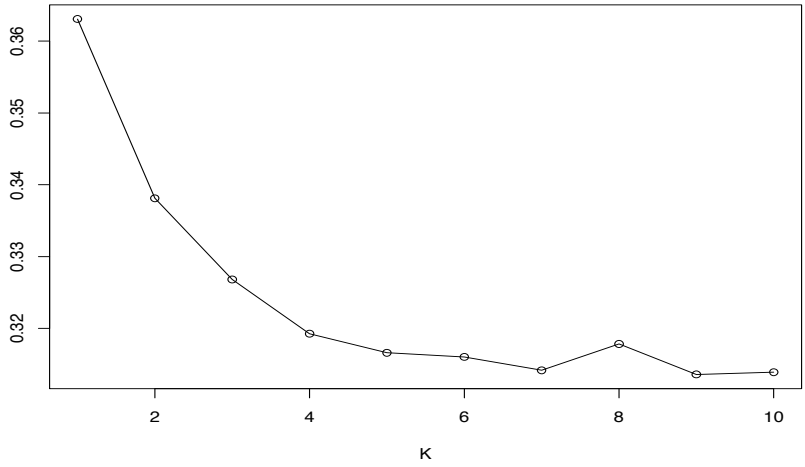
Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

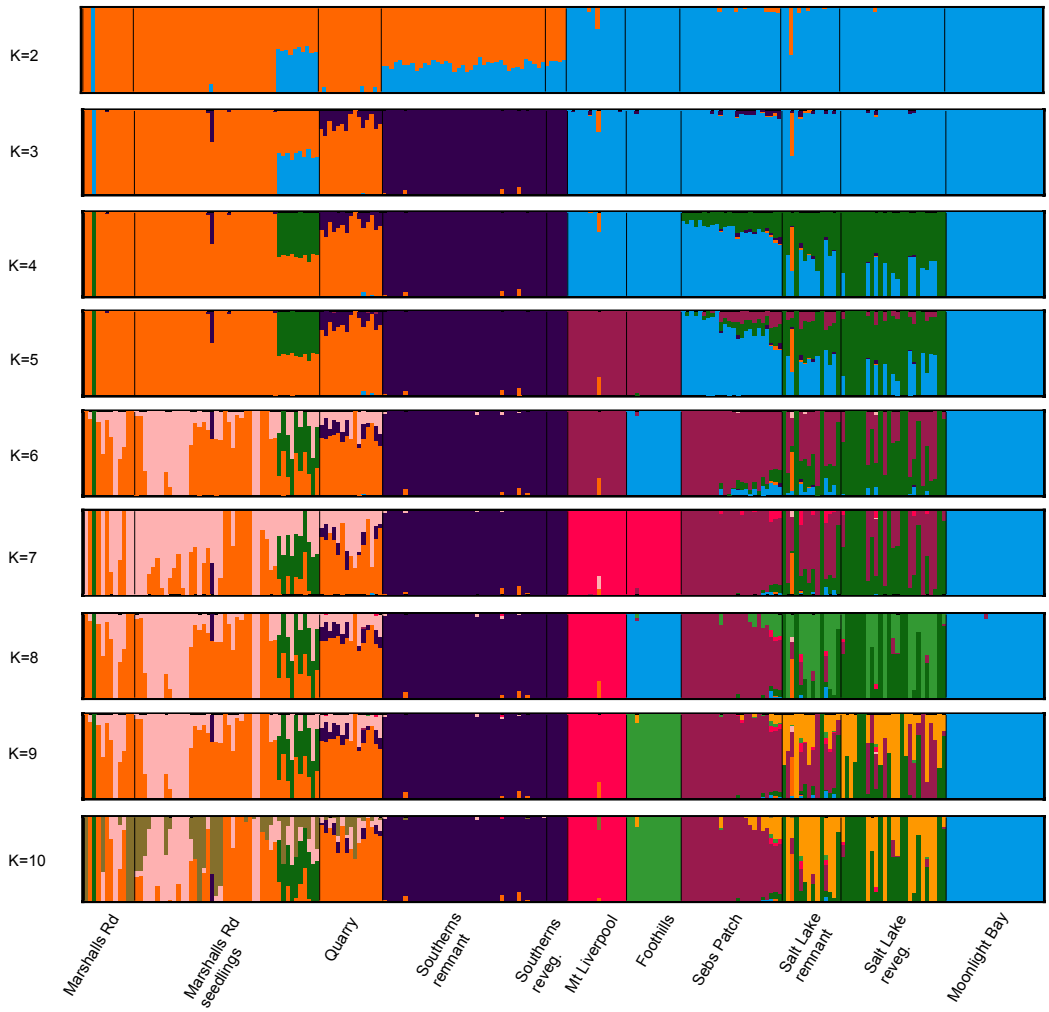
Table A1. Summary of collection stands, habitat type, census counts, sample size after filtering (*n*), and land use history of *Acacia whibleyana*.

Collection Stand	*Habitat Type*	*Census Count*	*n*	*Land Use History*
Marshalls Road	Roadside vegetation.	42 reveg.	12	Road reserve. Control burn in 2018 resulted in *A. whibleyana* plants emerging post fire.
Quarry	Post-disturbance regeneration in disused quarry, surrounded by remnant eucalypt woodland	100+ seedlings	44	Council-owned land. Quarried up to 1950's, stockpiling gravel then woodchips (to 1993).
		189 remnant	15	
Southerns	Remnant eucalypt woodland.	59 remnant	39	Private land. 40ha fenced to exclude stock in May 2012.
		8 reveg.	5	
Mount Liverpool	Remnant eucalypt woodland.	50 remnant	14	Private land subdivided for rural living & stock removed 2005.
Foothills Road	Roadside vegetation.	17 remnant	13	Road reserve, burnt in 2005 bushfires. *A. whibleyana* plants emerged post fire.
Sebs Patch	Remnant chenopod scrub along salt-scalded drainage line and salt lake.	107 remnant	24	Private land under different land management regimes (grazed by sheep intermittently, a small portion fenced & revegetated)
Salt Lake	Remnant chenopod scrub adjacent to salt lake.	137 remnant	14	Council-owned land of road intersection, fenced in 1995 to exclude vehicles and rabbits.
		26 reveg.	25	
Moonlight Bay	Remnant mallee woodland on salt lake lunettes.	557 remnant	23	Private land. Approx. one third of population fenced from stock. Remainder accessible to sheep on intermittent basis. Area around salt lake burnt in 2005 bushfires.

Note. Adapted from Faast, R.; Blyth, C.; Breed, M.F.; Packer, J.G. *Acacia Whibleyana: Part 1—Literature Synthesis to Inform Conservation of a Threatend Acacia*; Eyre Peninsula Natural Resources Management Board, 2019, Port Lincoln, South Australia, Australia.



(**a**)



(**b**)

Figure A1. (**a**) ADMIXTURE error results - the cv errors are plotted for K = 1–10, with the lowest cv error indicating greatest support at K = 7. (**b**) Bar plots represent individual genetic cluster assignment from ADMIXTURE results from K = 2 to K = 10.

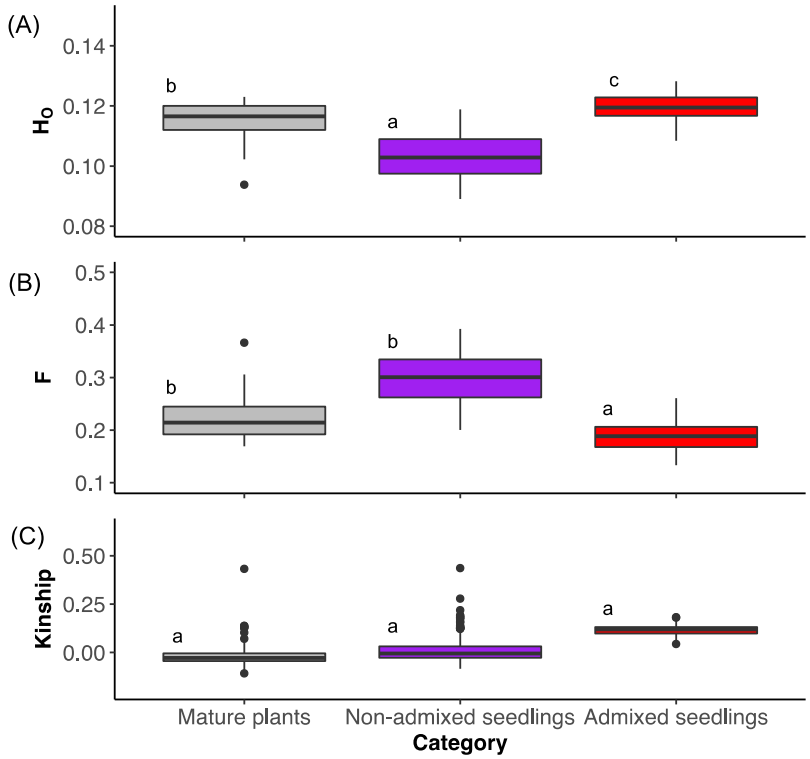


Figure A2. Boxplots of the Marshalls Rd mature plants, admixed seedlings, and non-admixed seedlings. (**A**) proportions of single nucleotide polymorphisms (SNPs) that are heterozygous, (**B**) inbreeding coefficient (F), and (**C**) the kinship values. Homogeneous subgroups are indicated by a, b, and c. Note the greater proportion of heterozygous SNPs in admixed seedlings, demonstrating that admixture between genetic clusters may lead to greater genetic diversity compared to non-admixed seedlings and the parent populations.

References

1. Reed, D.H.; Frankham, R. Correlation between fitness and genetic diversity. *Conserv. Biol.* **2003**, *17*, 230–237. [CrossRef]
2. Nicotra, A.B.; Beever, E.A.; Robertson, A.L.; Hofmann, G.E.; O'Leary, J. Assessing the components of adaptive capacity to improve conservation and management efforts under global change. *Conserv. Biol.* **2015**, *29*, 1268–1278. [CrossRef] [PubMed]
3. Leimu, R.; Mutikainen, P.; Koricheva, J.; Fischer, M. How general are positive relationships between plant population size, fitness and genetic variation? *J. Ecol.* **2006**, *94*, 942–952. [CrossRef]
4. Aguilar, R.; Quesada, M.; Ashworth, L.; Herrerías-Diego, Y.; Lobo, J. Genetic consequences of habitat fragmentation in plant populations: Susceptible signals in plant traits and methodological approaches. *Mol. Ecol.* **2008**, *17*, 5177–5188. [CrossRef]
5. Keller, L. Inbreeding effects in wild populations. *Trends Ecol. Evol.* **2002**, *17*, 230–241. [CrossRef]
6. Spielman, D.; Brook, B.W.; Frankham, R. Most species are not driven to extinction before genetic factors impact them. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15261–15264. [CrossRef]
7. Weeks, A.; Sgrò, C.M.; Young, A.G.; Frankham, R.; Mitchell, N.J.; Miller, K.A.; Byrne, M.; Coates, D.J.; Eldridge, M.D.B.; Sunnucks, P.; et al. Assessing the benefits and risks of translocations in changing environments: A genetic perspective. *Evol. Appl.* **2011**, *4*, 709–725. [CrossRef]
8. Barmentlo, H.; Meirmans, P.G.; Luijten, S.H.; Triest, L.; Oostermeijer, G. Outbreeding depression and breeding system evolution in small, remnant populations of Primula vulgaris: Consequences for genetic rescue. *Conserv. Genet.* **2017**, *19*, 545–554. [CrossRef]
9. Frankham, R.; Ballou, J.D.; Eldridge, M.D.B.; Lacy, R.C.; Ralls, K.; Dudash, M.R.; Fenster, C.B. Predicting the probability of outbreeding depression. *Conserv. Biol.* **2011**, *25*, 465–475. [CrossRef]

10. Ralls, K.; Dudash, M.R.; Eldridge, M.D.B.; Fenster, C.B.; Sunnucks, P.; Ballou, J.D.; Lacy, R.; Frankham, R. Call for a paradigm shift in the genetic management of fragmented populations. *Conserv. Lett.* **2017**, *11*, e12412. [CrossRef]
11. Pickup, M.; Field, D.L.; Rowell, D.M.; Young, A.G. Source population characteristics affect heterosis following genetic rescue of fragmented plant populations. *Proc. R. Soc. B Biol. Sci.* **2013**, *280*, 20122058. [CrossRef] [PubMed]
12. Hoffmann, A.; Sgrò, C.M. Climate change and evolutionary adaptation. *Nature* **2011**, *470*, 479–485. [CrossRef] [PubMed]
13. Bell, D.A.; Robinson, Z.L.; Funk, W.C.; Fitzpatrick, S.W.; Allendorf, F.W.; Tallmon, D.A.; Whiteley, A.R. The exciting potential and remaining uncertainties of genetic rescue. *Trends Ecol. Evol.* **2019**, *34*, 1070–1079. [CrossRef] [PubMed]
14. Williams, A.V.; Nevill, P.G.; Krauss, S.L. Next generation restoration genetics: Applications and opportunities. *Trends Plant Sci.* **2014**, *19*, 529–537. [CrossRef] [PubMed]
15. Carvalho, Y.G.S.; Vitorino, L.C.; De Souza, U.J.B.; Bessa, L.A. Recent trends in research on the genetic diversity of plants: Implications for conservation. *Diversity* **2019**, *11*, 62. [CrossRef]
16. Breed, M.F.; Harrison, P.A.; Blyth, C.; Byrne, M.; Gaget, V.; Gellie, N.J.C.; Groom, S.V.; Hodgson, R.; Mills, J.G.; Prowse, T.A.A.; et al. The potential of genomics for restoring ecosystems and biodiversity. *Nat. Rev. Genet.* **2019**, *20*, 615–628. [CrossRef]
17. Ottewell, K.; Bickerton, D.C.; Byrne, M.; Lowe, A.J. Bridging the gap: A genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. *Divers. Distrib.* **2015**, *22*, 174–188. [CrossRef]
18. Commonwealth of Australia. *Threatened Species Strategy—Year One Report*; Department of the Environment and Energy: Alice Springs, NT, Australia, 2016.
19. Jusaitis, M. Translocation trials confirm specific factors affecting the establishment of three endangered plant species. *Ecol. Manag. Restor.* **2005**, *6*, 61–67. [CrossRef]
20. Jusaitis, M.; Polomka, L. Weeds and propagule type influence translocation success in the endangered Whibley wattle, Acacia whibleyana (Leguminosae: Mimosoideae). *Ecol. Manag. Restor.* **2008**, *9*, 72–76. [CrossRef]
21. Faast, R.; Blyth, C.; Breed, M.F.; Packer, J.G. *Acacia Whibleyana: Part 1—Literature Synthesis to Inform Conservation of a Threatend Acacia*; Eyre Peninsula Natural Resources Management Board: Port Lincoln, South Australia, 2019.
22. Adams, M.; Attiwill, P. Role of *Acacia* spp. in nutrient balance and cycling in regenerating *Eucalyptus regnans* F. Muell. forests. I. Temporal changes in biomass and nutrient content. *Aust. J. Bot.* **1984**, *32*, 205–215. [CrossRef]
23. Jusaitis, M.; Sorensen, B. *Conservation Biology of Acacia Whibleyana*; Black Hill Flora Centre, Botanic Gardens of Adelaide: Adelaide, Australia, 1998.
24. Peterson, B.K.; Weber, J.N.; Kay, E.H.; Fisher, H.S.; Hoekstra, H.E. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* **2012**, *7*, e37135. [CrossRef] [PubMed]
25. Catchen, J.M.; Amores, A.; Hohenlohe, P.; Cresko, W.A.; Postlethwait, J.H. Stacks: Building and genotyping loci de novo from short-read sequences. *G3 Genes Genomes Genet.* **2011**, *1*, 171–182. [CrossRef] [PubMed]
26. Catchen, J.; Hohenlohe, P.A.; Bassham, S.; Amores, A.; Cresko, W.A. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **2013**, *22*, 3124–3140. [CrossRef] [PubMed]
27. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef] [PubMed]
28. Purcell, S.M.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.B.; Sklar, P.; De Bakker, P.I.W.; Daly, M.J.; et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
29. Meirmans, P.G. genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Mol. Ecol. Resour.* **2020**. [CrossRef]
30. Shafer, A.B.A.; Peart, C.R.; Tusso, S.; Maayan, I.; Brelsford, A.; Wheat, C.W.; Wolf, J.B.W. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods Ecol. Evol.* **2016**, *8*, 907–917. [CrossRef]
31. Jombart, T.; Devillard, S.; Balloux, F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet.* **2010**, *11*, 94. [CrossRef]

32. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [CrossRef]
33. Rosenberg, N.A. Distruct: A program for the graphical display of population structure. *Mol. Ecol. Notes* **2003**, *4*, 137–138. [CrossRef]
34. Kopelman, N.M.; Mayzel, J.; Jakobsson, M.; Rosenberg, N.A.; Mayrose, V.B. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **2015**, *15*, 1179–1191. [CrossRef] [PubMed]
35. Wickham, H. ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **2011**, *3*, 180–185. [CrossRef]
36. Ottewell, K.; Bickerton, D.; Lowe, A.J. Can a seed bank provide demographic and genetic rescue in a declining population of the endangered shrub *Acacia pinguifolia*? *Conserv. Genet.* **2010**, *12*, 669–678. [CrossRef]
37. Stone, G.N.; Raine, N.E.; Prescott, M.; Willmer, P.G. Pollination ecology of acacias (Fabaceae, Mimosoideae). *Aust. Syst. Bot.* **2003**, *16*, 103–118. [CrossRef]
38. Breed, M.F.; Ottewell, K.; Gardner, M.G.; Marklund, M.H.K.; E Dormontt, E.; Lowe, A.J. Mating patterns and pollinator mobility are critical traits in forest fragmentation genetics. *Heredity* **2013**, *115*, 108–114. [CrossRef]
39. Jusaitis, M.; Sorensen, B.; Polomka, L. Population structure, flowering and seed production in the endangered whibley wattle, *Acacia whibleyana* (Leguminosae: Mimosoideae). *Open Conserv. Biol. J.* **2009**, *3*, 4–13. [CrossRef]
40. Petit, R.J.; Duminil, J.; Fineschi, S.; Hampe, A.; Salvini, D.; Vendramin, G.G. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol. Ecol.* **2004**, *14*, 689–701. [CrossRef]
41. Borrell, J.S.; Wang, N.; Nichols, R.A.; Buggs, R. Genetic diversity maintained among fragmented populations of a tree undergoing range contraction. *Heredity* **2018**, *121*, 304–318. [CrossRef]
42. De Kort, H.; Vandepitte, K.; Bruun, H.H.; Closset-Kopp, D.; Honnay, O.; Mergeay, J. Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree speciesAlnus glutinosa. *Mol. Ecol.* **2014**, *23*, 4709–4721. [CrossRef]
43. Johnson, J.; Gaddis, K.D.; Cairns, D.M.; Konganti, K.; Krutovsky, K.V. Landscape genomic insights into the historic migration of mountain hemlock in response to Holocene climate change. *Am. J. Bot.* **2017**, *104*, 439–450. [CrossRef]
44. Pais, A.; Whetten, R.W.; Xiang, Q. (Jenny) Ecological genomics of local adaptation in *Cornus florida* L. by genotyping by sequencing. *Ecol. Evol.* **2016**, *7*, 441–465. [CrossRef] [PubMed]
45. He, T.; D'Agui, H.; Lim, S.L.; Enright, N.J.; Luo, Y. Evolutionary potential and adaptation of *Banksia attenuata* (Proteaceae) to climate and fire regime in southwestern Australia, a global biodiversity hotspot. *Sci. Rep.* **2016**, *6*, 26315. [CrossRef] [PubMed]
46. Kenrick, J.; Knox, R.B. Quantitative analysis of self-incompatibility in trees of seven species of *Acacia*. *J. Hered.* **1989**, *80*, 240–245. [CrossRef]
47. Kenrick, J.; Kaul, V.; Williams, E.G. Self-incompatibility in *Acacia retinodes*: Site of pollen-tube arrest is the nucellus. *Planta* **1986**, *169*, 245–250. [CrossRef] [PubMed]
48. Millar, M.A.; Byrne, M.; Nuberg, I.; Sedgley, M. High outcrossing and random pollen dispersal in a planted stand of *Acacia saligna* subsp. *saligna* revealed by paternity analysis using microsatellites. *Tree Genet. Genomes* **2007**, *4*, 367–377. [CrossRef]
49. Lesica, P.; Allendorf, F.W. When are peripheral populations valuable for conservation? *Conserv. Biol.* **1995**, *9*, 753–760. [CrossRef]
50. Coates, D.J.; Tischler, G.; McComb, J.A. Genetic variation and the mating system in the rare *Acacia sciophanes* compared with its common sister species *Acacia anfractuosa* (Mimosaceae). *Conserv. Genet.* **2006**, *7*, 931–944. [CrossRef]
51. Nano, C.; Nano, T.; Gibson, J.; Pavey, C. *Recovery Action Implementation for Threatened Arid Acacias: Distribution, Monitoring and Indigenous Ecological Knowledge of A. Peuce, A. Undoolyana, A. Pickardii and A. Latzii*; Department of Natural Resources, Environment, Arts & Sport Northern Territory Government: Alice Springs, Australia, 2008.

Statement of Authorship

Title of Paper	Population genomics for delineating rangewide seed zones for restoration of a keystone arid-zone shrub
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ■ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Blyth, C., Breed, M.F., Hodgson R. J., Gellie, N, J. C., Lowe, A.J, McDonald, J., Christmas, M.J.,. Population genomics for delineating rangewide seed zones for restoration of a keystone arid-zone shrub. In Prep for submission to Restoration Ecology

Principal Author

Name of Principal Author (Candidate)	Colette Blyth		
Contribution to the Paper	Helped develop the ideas and methods, conducted fieldwork, analysed the data and wrote the manuscript.		
Overall percentage (%)	60		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr Martin Breed		
Contribution to the Paper	Helped secure funding, helped develop the ideas and methods, contributed to draft revisions of manuscript.		
Signature		Date	2/7/2021
Name of Co-Author	Riley Hodgson		
Contribution to the Paper	Helped with fieldwork, contributed to draft revisions of manuscript.		

Signature		Date	2/7/21
Name of Co-Author	Dr Nicolas Gellie		
Contribution to the Paper	Helped develop the ideas and methods, helped with fieldwork, contributed to draft revisions of manuscript.		
Signature		Date	2/7/2021
Name of Co-Author	Professor Andy Lowe		
Contribution to the Paper	Helped secure funding, helped develop the ideas and methods, contributed to draft revisions of manuscript.		
Signature		Date	5/7/2021
Name of Co-Author	Dr John McDonald		
Contribution to the Paper	Helped with fieldwork, helped develop the ideas and methods, contributed to draft revisions of manuscript.		
Signature		Date	13/7/2021
Name of Co-Author	Dr Matthew Christmas		
Contribution to the Paper	Helped analyse the data and contributed to draft revisions of manuscript.		
Signature		Date	8/7/2021

Chapter 3. Population genomics for delineating rangewide seed zones for restoration of a keystone arid-zone shrub

Colette Blyth[1], Martin F. Breed[2], Riley J. Hodgson[2], Nicholas J. C. Gellie[1,3],

Andrew J. Lowe[1], John McDonald[1], Matthew J. Christmas[4]

[1]School of Biological Sciences, University of Adelaide, Adelaide 5005, Australia

[2]College of Science and Engineering, Flinders University, Bedford Park 5042, Australia

[3]Terrestrial Ecosystem Research Network (TERN), University of Adelaide, Adelaide 5005, Australia

[4]Department of Medical Biochemistry and Microbiology, Uppsala University, 751 23 Uppsala, Sweden

Abstract

The success of landscape restoration projects relies heavily on sourcing large quantities of highly viable seed, and a key decision is where in the landscape to source this seed. Traditionally seed is preferentially collected local to restoration sites because this is seen to favour local adaptation but, with climate change, collecting seed from a wider source may increase the adaptive potential of restored populations. However, broader collection zones come with risks of introducing inappropriate genetic lineages, that could cause maladaptation or outbreeding depression. Therefore, accurately genetically delimiting seed zones within which seed can be appropriately sourced is important for restoration projects. Population genomics can be used to accurately define seed zones, but its application to date has been limited. In this study, we define range-wide seed zones for the shrub *Maireana sedifolia* from a genomic dataset spanning >1500 km east-west across arid southern Australia. We identified two distinct genetic clusters and used these to define preliminary seed transfer zones. Overall, we found low population structure, reflecting species life (long lived, wind-dispersed) and biogeographical history. We then analysed variation in the environment between these genetic clusters, finding significant differences between the environments of both clusters, and between both these clusters and some genetically admixed collection localities We discuss our findings in a biogeographic context of the Australian arid zone, provide seed sourcing recommendations, and present options to further explore the adaptive significance of populations in these seed zones.

Introduction

Approximately one third of all arable land is now degraded, driven largely by broadscale clearing of native vegetation (Gibbs and Salmon, 2015, Nkonya et al., 2016). Restoration is employed as a direct response to land degradation, with the primary goal being to aid the recovery of lost biodiversity and/or ecosystem functions (SER, 2004). There has been enormous investment in restoration projects in recent years (Breed et al., 2013, Feng et al., 2013, van Woerdon, 2014) and a key step in this process is the selection of suitable plant genetic resources for revegetation (Clewell and Rieger, 1997, Crouzeilles et al., 2016).

Revegetation success rates a number of limitations (Crouzeilles et al., 2016, Godefroid et al., 2011), and poor seed quality can be a key restraint (Breed et al., 2013, Broadhurst et al., 2008, Godefroid et al., 2011). If revegetation projects are to proceed and be successful, access to large quantities of high-quality seed is vital (Breed et al., 2013, Broadhurst et al., 2008, Godefroid et al., 2011). Typically, practitioners have collected seed from local plant populations, which have been assumed to be well adapted to the target restoration site (Breed et al., 2013, Broadhurst et al., 2008). Selecting seed locally also minimises risks associated with introducing novel genotypes to the restoration site that could result in maladaptation and outbreeding depression (Hufford and Mazer, 2003, Krauss et al., 2013, McKay et al., 2005). However, these guidelines do not account for the long-term impacts of habitat fragmentation and climate change, which can reduce connectivity and lead to a loss of genetic diversity that reduces the likelihood of populations adapting to future climate conditions (Bijlsma and Loeschcke, 2011, Frankham, 2005). Furthermore, climate change can shift adaptive landscapes faster than plants can respond, causing potential maladaptation of the local provenance (Christmas et al., 2015, Gellie et al, 2016).

Since restoration is generally carried out in fragmented habitats with populations vulnerable to inbreeding and erosion of genetic diversity, provenance selection is a crucial consideration. As such, restoration practitioners are now employing strategies that focus on establishing populations that are resilient to future climate change. By sourcing seed that elevates the genetic diversity of target populations, practitioners can increase the adaptive potential of these populations and introduce phenotypes that are better adapted to predicted future climates (Breed et al., 2016, Breed et al., 2013, Broadhurst et al., 2008, Prober et al., 2015). For example, climate-adjusted provenancing can be used to introduce genotypes from more arid areas into areas where aridity is predicted to increase, thus improving the survival chances of populations under projected future climate conditions (Prober et al., 2015).

Despite the benefits of increasing genetic diversity for restoration programs through moving away from local only seed sourcing, a consensus is required on how to safely do. One option is to define seed zones by employing a genomics toolkit (Breed et al., 2019). Seed zones define boundaries for the collection of seed. Within the boundary, it is considered safe to transfer seed with minimal risk of introducing maladaptated genotypes. Seed zones can be delineated using ecoregions to predict likely local adaptation to environment (e.g., Pike et al. 2020), or they can be guided by using genetic data to detect breaks in gene flow (e.g., Jørgensen et al., 2016) or with more recent genomic advances, local adaptation (Breed et al., 2019). Increasing affordability of next generation sequencing (NGS) technologies over the last decade has facilitated research into the application of genetics to plant species management (Blyth et al., 2021, Blyth et al., 2020, Broadhurst et al., 2021, De Kort et al., 2014, Rossetto et al., 2020). These studies have successfully used genetic analyses to reveal population genetic diversity and structure, identified genetic barriers, and in some cases, highlighted major selection pressures to the populations. Such information can be used to guide restoration and identify genetically suitable seed sources for restoration projects, increasing a population's chances of adapting to future climatic

changes, whilst minimising the risks of outbreeding depression associated with moving seed between disparate populations with large genetic breaks (Breed et al., 2019, Carvalho et al., 2019, Williams et al., 2014).

Maireana sedifolia F. Muell., commonly referred to as pearl bluebush, is a long lived, widespread, and often dominant species of chenopod shrub found in arid regions of southern Australia (Figure 1). It has a mostly continuous range spanning >1500 km and is distributed across known biogeographical barriers to gene flow for other plant species in this region, including the Flinders Ranges (Christmas et al., 2017), and the 'Nullarbor Plain Barrier' (Crisp and Cook, 2007) (Figure A1), suggesting that the species is adapted to and can disperse over, a wide range of environments. *Maireana sedifolia* is known to tolerate arid climates and calcareous soils, which many other plant species cannot, making it an important species in arid zone ecosystems (McDonald et al., 2017) and therefore critical to many restoration projects in this region. However, *M. sedifolia* is notoriously difficult to use in revegetation (Butler, 1985; McDonald et al., 2017). The species sets seed unpredictably and natural establishment is rare, with a recent literature synthesis estimating desirable conditions (sequential years of sufficient rainfall alongside a disturbance event) only likely to occur every 25 – 50 years (McDonald et al., 2017). There are many large mining sites found within the range of *M. sedifolia*, so it is often a target species for restoration and rehabilitation projects. However, practitioners have often struggled to access sufficient local seed [pers comms] for the species. Therefore, defining seed-sourcing boundaries based on population genetic diversity and divergence will provide practitioners with the information they need to maximise seed sourcing whilst mitigating the risks of the adverse effects of inbreeding and outbreeding depression.

Here, we carried out a population genomics study of 36 populations of *M. sedifolia* from across the species' range and measured levels of genetic diversity and divergence. Furthermore, we analysed environmental variation across the species' range and tested for associations between genotypic and environmental variation

with the aim of identifying significant environmental selection pressures. Lastly, we used our findings to delineate range wide seed zones to inform restoration practice and provide recommendations for future research into this recalcitrant seeding species.

Methods

Study species and sample collection

Maireana sedifolia is a perennial dioecious shrub in the sub-family Camphorosmoideae, of the family Chenopodiaceae and has an estimated lifespan of 150-300 years (Crisp, 1978; Cunningham et al., 1981). The species is found in the arid zone across Western Australia, South Australia, New South Wales and Victoria. Populations with the highest density tend to be found on soils which are calcareous and easily permeated by rainwater (Cunningham et al., 1981; McDonald et al., 2017; Specht 1972). Both pollen and seed are likely to be wind dispersed but flowering and seeding occurs unpredictably, suggesting that populations persist, at least in part, due to their long lifespan (Blackwell and Powell, 1981; McDonald et al., 2017). There is no record in the literature of the species reproducing clonally. Levels of ploidy within the species are currently unknown, however a recent study of the related Australian chenopod, *Sceloraena napiformis* found the species to be diploid (Amor et al. 2020).

We conducted range-wide sampling between 2014 and 2016 (Figure 1., Table A1). We sampled 6-8 individuals at each site across 36 collection localities. We avoided sampling nearest neighbours to minimise the likelihood of genotyping closely related individuals. Young leaves were targeted to maximise DNA quality and placed in gauze bags and desiccated on silica gel to dry until DNA extraction.

DNA extraction, sequencing, and SNP calling

DNA was extracted with the Machery-Nagel Nucleospin Plant II Kit at the Australian Genome Research Facility (AGRF; Adelaide, Australia). Double digest

restriction-associated DNA sequencing (ddRADseq; (Peterson et al., 2012) was used to generate genome-wide SNP markers for all samples. Full details of the method can be found in Peterson et al. (2012). Briefly, ddRAD library preparation protocol involved the following steps: (1) DNA digestion with two restriction enzymes (EcoRI and MseI); (2) ligation of barcoded adaptors, specific to each sample, to restriction site overhangs; (3) a wide size selection of pooled digested-ligated fragments 280-375 bp (95 bp range) using Blue Pippin (Sage Science, Beverly, MA, USA) (Wide); (4) amplification of library via PCR for 12 cycles using indexed primers. Libraries were assessed by gel electrophoresis (Agilent D1000 ScreenTape Assay), quantified by qPCR (KAPA Library Quantification Kits for Illumina) and then sequenced with 150 bp single reads on the NextSeq 500 system using NextSeq 500/550 Mid Output Kit v2 (150 cycles) reagents. This resulted in 136,516,593 sequences across the 286 samples.

Following sequencing, reads were processed with the STACKS pipeline (Catchen et al., 2011, Catchen et al., 2013) at AGRF (Melbourne, Australia). Briefly, reads were deconvoluted by inline barcodes, checked for read quality and trimmed. After trimming, stacks of similar reads were created for each sample individually, these read stacks are also known as tags. Tags which appear across all samples were collated (catalogue tags), and genotypes were then allocated to the common polymorphic sites.

The collated SNPs across all individuals were then filtered with the following settings: minimum number of reads required at a stack to call a homozygous genotype = 5; minor allele frequency, below which a stack is called a homozygote = 0.05; minimum minor allele frequency to call a heterozygote = 0.1 (above 0.05 but below 0.1, a stack is called 'unknown'). This resulted in a total of 556,907 SNPs, which were then output as a single vcf file for further filtering and analysis.

SNP filtering

We used VCFtools (Danecek et al., 2011) to remove SNPs with a minimum depth of coverage of <10 reads per individual and a minor allele frequency of >0.1. PLINK (Purcell et al., 2007). SNPs with >25% missing and individuals were removed if they had >25% missing SNP calls. SNPs in linkage disequilibrium were identified with the --indep-pairwise command in PLINK — one SNP in a pair of linked SNPs (r^2 > 0.8) was pruned from the data set. A negative value of *F* (inbreeding coefficient) represents higher than expected heterozygosity under Hardy-Weinberg equilibrium and may be indicative of paralogous reads being stacked together in the STACKS pipeline (Shafer et al., 2016). Therefore, *F* values were calculated for each SNP in GENODIVE v. 2.0b27 (Meirmans and Van Tienderen, 2004) and any with significantly negative *F* were removed using VCFtools (we ran permutation tests with 10,000 permutations and a significance threshold of $p < 0.05$). These filtering steps resulted in a final set of 5,672 SNPs across 270 individuals for downstream analysis.

Genetic structure analysis

The filtered SNP set was used for subsequent analyses to investigate population genetic structure and diversity and to explore whether variation in the data was found to associate with environmental variables. To assess genetic structure, we used two different clustering analyses to identify the most likely number of genetic clusters (K) in our dataset. Firstly, we used the non-model based Discriminant Analysis of Principal Components (DAPC) using the R package adegenet (Jombart et al., 2010) to calculate likely for K = 1 to 10. Then, the model-based genetic clustering algorithm ADMIXTURE (Alexander et al. 2009) was also used to estimate the most likely number of clusters (K) in our dataset. ADMIXTURE includes a cross-validation procedure for identifying the value of K for which the model has best predictive accuracy. ADMIXTURE was also run for K values 1 to 10 and the most likely value of K was assessed by comparing the cross-validation errors (cv errors) between runs,

with the lowest cv error indicating greatest support. Barplots of individual assignment to clusters for the most supported value of K were created in DISTRUCT (Rosenberg, 2004). As several populations contained a high degree of admixture, we further explored the data to ascertain whether this was true admixture or whether the resolution of the data did not allow confident assignments of the individuals to the genetic groups identified. As the confidence intervals of the highly admixed populations did not overlap with the assignment of 1 or 0, we concluded that the populations were indeed intermediate (Table A2) Then, to calculate genetic differentiation between the collection localities, the R package StamPP (Pembleton et al., 2013) was used to calculated pairwise F_{ST} between all collection localities. The output was exported to Splitstree v.4 to create a tree from the F_{ST} values (Huson, 1998).

These methods all suggested that the most likely number of genetic clusters in the data is two, with most individuals having a high proportion of assignment to one of the two clusters. A subset of individuals displayed a large proportion of admixture between the two clusters and we therefore grouped together all individuals with <80% to a single genetic cluster as a separate group to further explore the effect of admixture.

Before calculating our genetic diversity statistics, we accounted for sampling bias between our genetic clusters. First, we selected a random subset of individuals from the larger genetic cluster, so we had a sample number which matched the smaller genetic cluster. We then filtered the subset and smaller genetic cluster combined for a minor allele frequency of >0.1. PLINK (Purcell et al., 2007). Any SNPs filtered out of this dataset were then filtered out of the larger dataset containing all individuals, leaving 5,363 SNPs. Using this dataset, we conducted a nested Analysis of Molecular Variance (AMOVA) using GENODIVE v3.2 (Meirmans, 2020) to explore genetic variation within individuals, among collection localities nested within genetic clusters, and among genetic clusters.

Genetic Diversity and Inbreeding Analysis

We estimated the inbreeding coefficient (*F*) and observed (H_O) and expected (H_E) heterozygosity in GENODIVE v3.2 (Meirmans, 2020). A one-way ANOVA and a Tukey post-hoc pairwise comparisons were used to assess differences in *F*, H_O and H_E between the genetic clusters using R v.4.0.4 (Team, 2020). We again categorised collection localities with >20% genetic admixture as a separate group.

Environmental variation analysis

A set of 44 environmental (e.g., climate, soil, and landscape) variables were sourced from stacks of 9-s resolution (i.e., approximately 250 m pixels) rasters (Gallant et al., 2018, Grundy et al., 2015, Harwood et al., 2016). Environmental values were extracted for each raster at the coordinates of collection localities of *M. sedifolia* using the RASTER analytical package in R (Hijmans, 2021). Collinearity amongst these environmental variables was assessed using the 'cor' function in R v.4.0.4 using a threshold of $r < 0.3$. (R Core Team, 2020) and a subset of variables with low collinearity was selected. This included aridity index (proportion), soil pH, precipitation (mm), temperature range (°C), soil phosphorus (%), soil nitrogen (%), soil bulk density (g/cm^3), soil carbon (%), minimum temperature (°C), and soil silt (%). A principal component analysis (PCA) was performed using the 'prcomp' function in R v.4.0.4 (R Core Team, 2020) on the environmental variables at all collection localities to assess differences in environments among western, eastern and admixed localities. One-way ANOVA and Tukey post-hoc pairwise comparisons were used to assess differences in principal component axes 1 (PC1) and 2 (PC2) between the western and eastern genetic clusters and collection localities with >20% genetic admixture R v.4.0.4 (Team, 2020).

Redundancy Analysis on genetic variation

To assess the degree to which the environment explains observed genetic variation, a partial redundancy analysis (RDA) of allele frequencies of all SNPs in all collection localities was performed. We used the same set of environmental variables as in the environmental PCA analysis. This analysis assessed how much of the genetic variation among our collection localities can be explained by spatial and environmental factors using the VEGAN analytical package v2.5-6 in R (Dixon, 2003). First, spatial coordinates for each sampling location were centered and converted into third degree polynomials. An RDA of allele frequencies against all spatial polynomials was then performed using the 'rda' function. The 'ordistep' function was then used to carry out forward stepwise model building for the RDA, adding in one polynomial to the model at a time and assessing significance using permutation tests. This revealed x, y, xy^2, x^2 and x^3 to be significant contributors to the model and as such only these variables were retained for the RDA. The variance between the selected spatial and all environmental variables was assessed using the 'varpart' function and significance of the partitioning was assessed using an analysis of variance (ANOVA)-like permutation test for RDA with the 'anova.cca' function. We tested whether any of the genetic variance could be significantly explained by environmental variables by running an RDA on the model: allele frequencies against environmental variables conditioned on spatial variables.

Results

Genetic structure analysis

We identified two main genetic clusters across our samples but weak genetic structure overall. We observed a clear genetic split in the species' geographic range with an east/west divide and some admixture in between the two genetic clusters. This was supported by results from the ADMIXTURE, DAPC and F_{ST} analyses, with all analyses yielding similar results (Figures 1, 2, A3, A3). Our AMOVA analysis revealed

that most (80.9%) of the genetic variance was distributed within individuals and a significant proportion (14.9%) of variance was within individuals nested in collection localities ($p < 0.001$; Table 1). Both variation among collection localities nested within genetic cluster and variation among the genetic groups were also significant (1.8%, $p < 0.001$ and 2.5%, $p < 0.001$ respectively). Pairwise genetic differentiation (F_{ST}) was low between collection localities across the whole dataset ($F_{ST} = 0.00 – 0.09$; Table A3). We do see slightly increasing differentiation with increasing geographic distance between localities (i.e., some effect of isolation by distance). The neighbour-joining tree based on F_{ST} (Figure 2) reflects the same east/west divide as seen in the ADMIXTURE and DAPC analysis, with populations with >20% admixture positioned in between the two clusters.

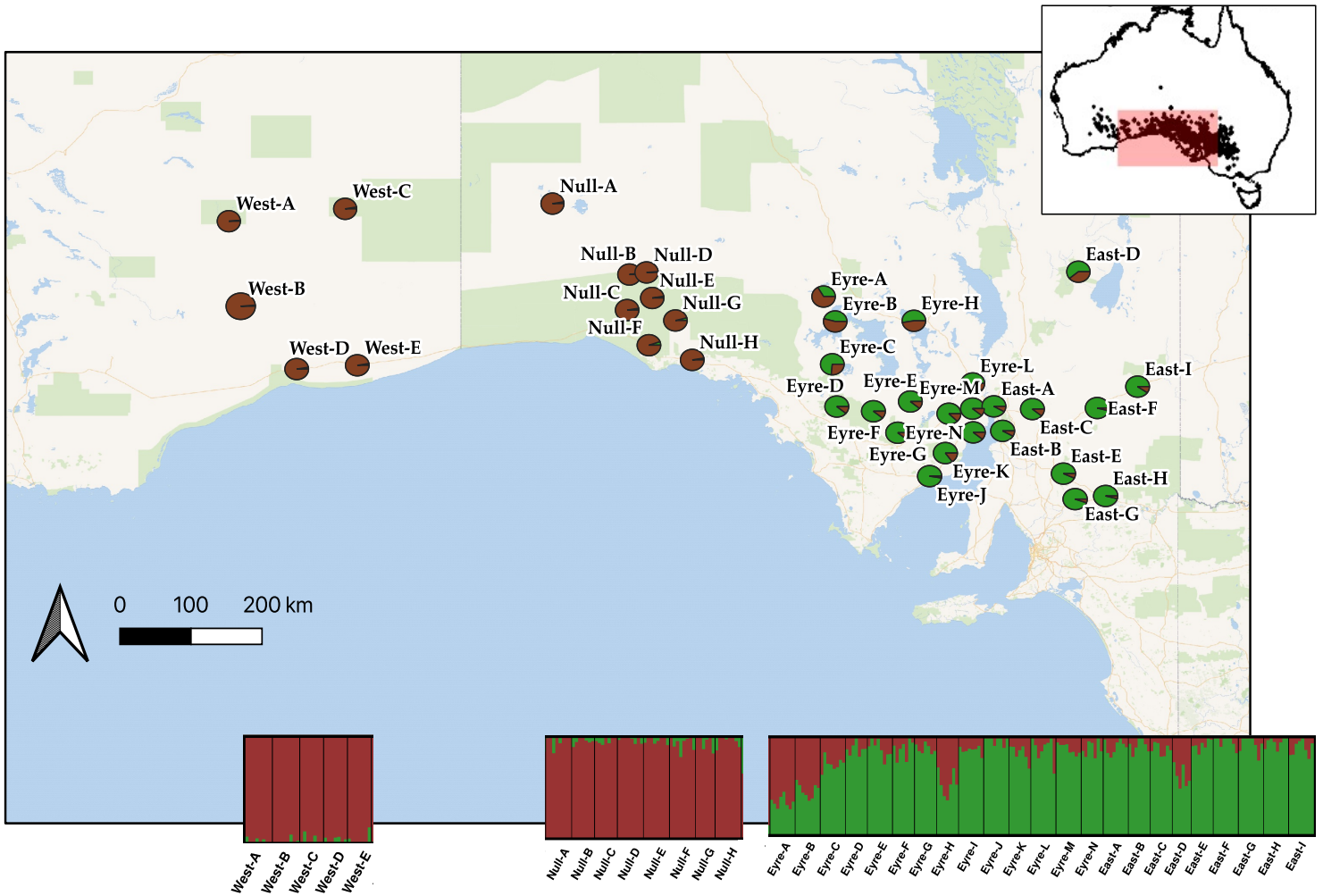


Figure 1. Individual genetic cluster assignment based on ADMIXTURE results. Pie charts at each sampling site show the overall proportion of assignment to the two genetic clusters, with a west/east divide. Individual assignments are represented in the bar plots at the bottom of the map. The black dots on the Australian continent map (inset) represent all *M. sedifolia* occurrence records on the Atlas of Living Australia (Atlas of Living Australia occurrence download at http://www.ala.org.au. Accessed 10 February 2017).

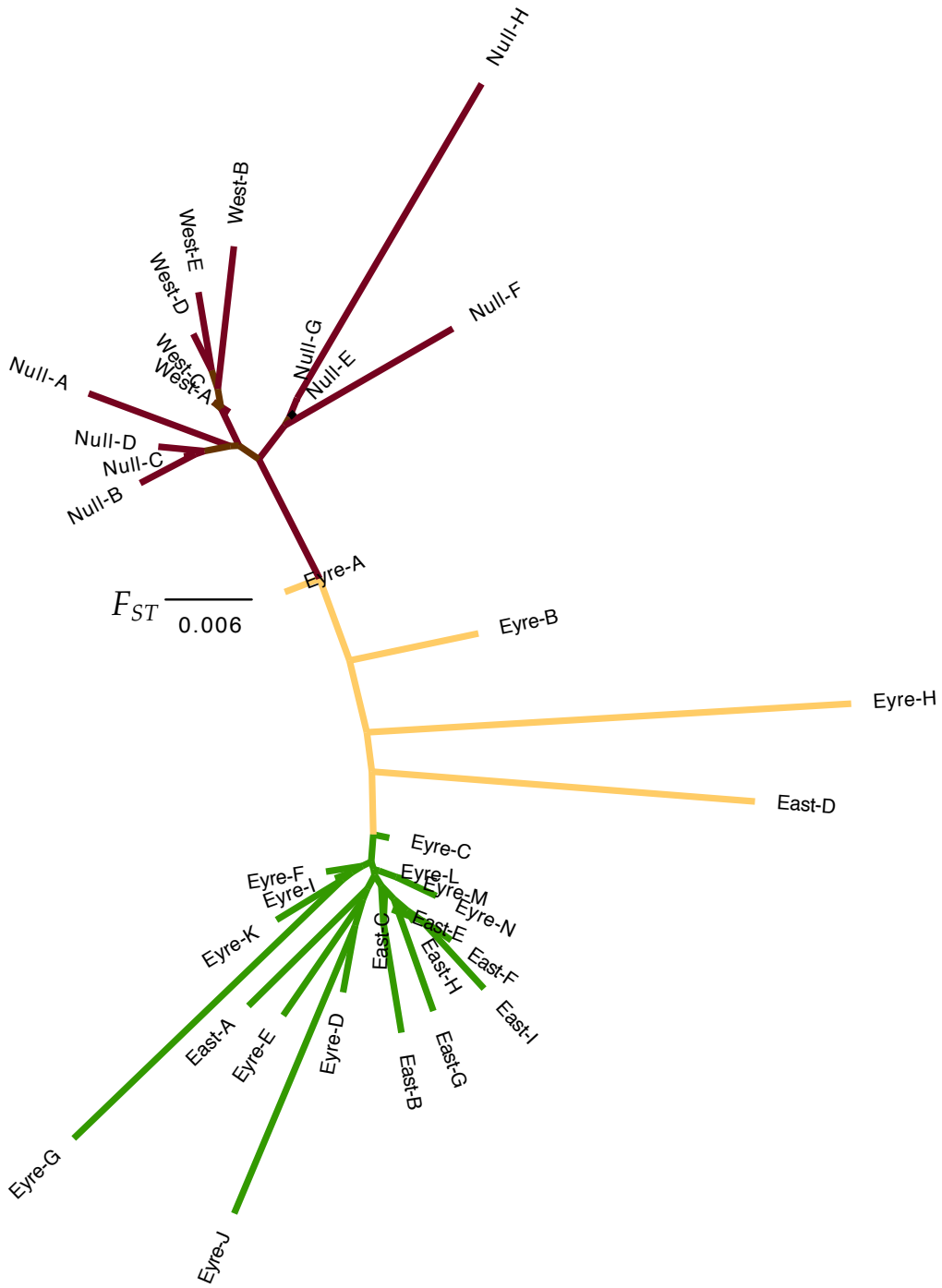


Figure 2. Neighbour joining tree of pairwise genetic differentiation (F_{ST}) between collection localities. Colours represent genetic cluster assignment identified in the ADMIXTURE analysis and also highlights collection localities with >20% admixture (brown: western genetic cluster, yellow: admixed collection localities, green: eastern genetic cluster).

Table 1. Nested analysis of molecular variance (AMOVA) analysing the distribution of genetic diversity across all collection localities nested within genetic cluster. Including standard deviations (SD; obtained through jackknifing over loci) and 95% confidence intervals (c.i. 95%; obtained through bootstrapping over loci) of F statistics.

Source of variation	Nested in	%var	F	Std.Dev.	c.i.2.5%	c.i.97.5%	*p*
Within individual	--	80.9	0.19	0.003	0.185	0.198	--
Among individuals	Collection locality	14.9	0.15	0.003	0.149	0.161	<0.001
Among collection localities	Genetic cluster	1.8	0.02	0.001	0.018	0.021	<0.001
Among genetic cluster	--	2.5	0.02	0.001	0.024	0.026	<0.001

Environmental variation

Our principal component analysis (PCA) of differences in environment at each collection locality found PC1 to largely represent variation in aridity, precipitation, temperature range, soil phosphorus, soil nitrogen, soil density, and pH. Principal component 2 largely reflects variation in soil carbon, minimum temperature, and soil silt (Figure 3). The ANOVA and Tukey post-hoc pairwise comparison test revealed that PC1 was significantly different between admixed and eastern cluster collection sites ($p = 0.008$; Table A4) but neither differed significantly from the western genetic cluster collection sites. For PC2, we found a significant difference between the eastern and western genetic cluster collection sites ($p < 0.001$; Table A4), but neither differed from the admixed sites on this axis.

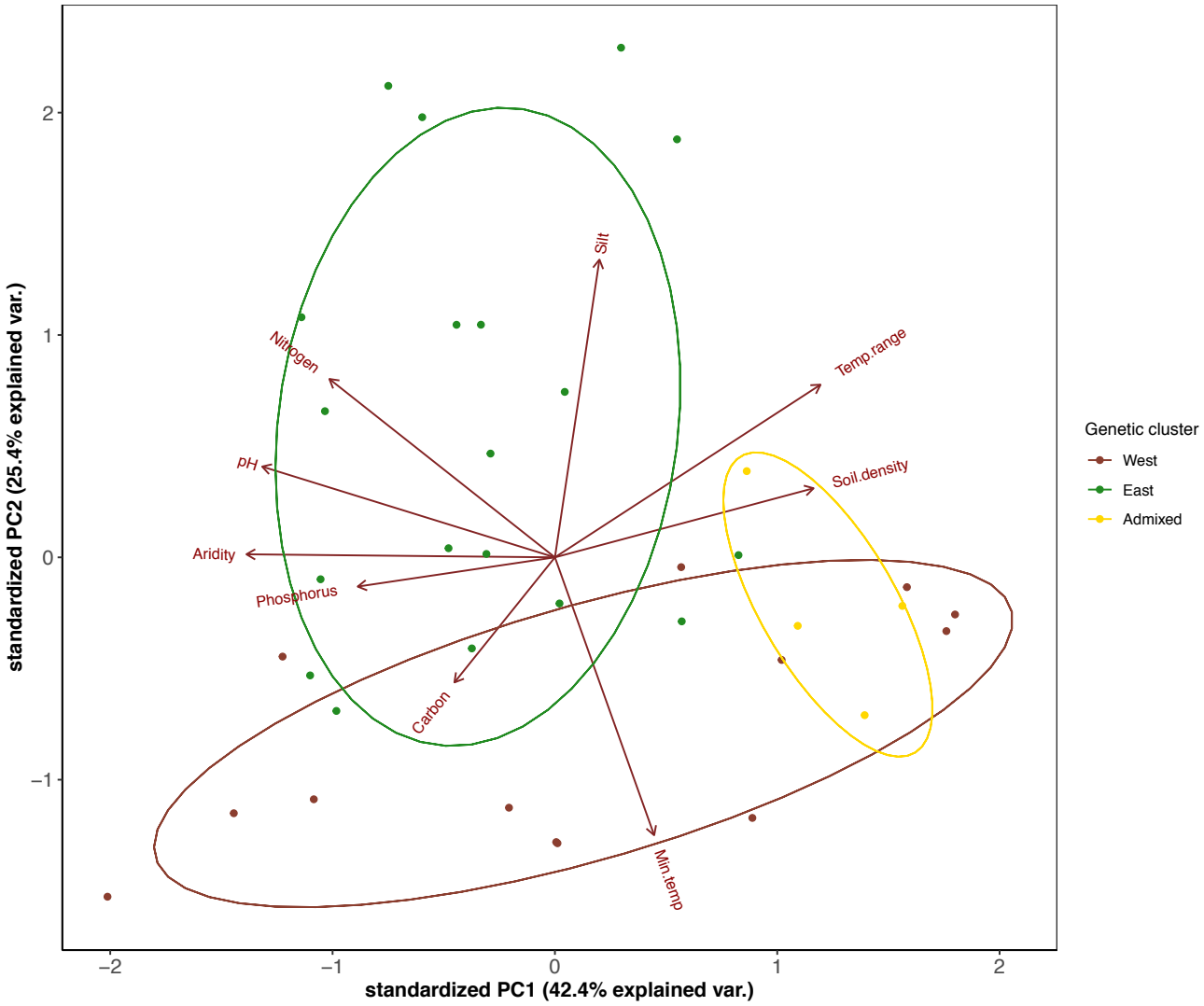


Figure 3. PCA of environmental variation across collection localities. Colours represent genetic cluster assignment identified in the ADMIXTURE analysis and highlights collection localities with >20% admixture (brown: western genetic cluster, yellow: admixed collection localities, green: eastern genetic cluster). The ellipses represent the core area added by the 68% confidence interval.

Redundancy analysis

The redundancy analysis revealed that spatial distribution significantly explained 14% of the variance in allele frequencies (ANOVA, F= 1.64, p = 0.001). The partial RDA, where we assessed the amount of variance explained by environment conditioned for that explained by spatial distribution, revealed that only 1% of the variance was explained by the environmental variables included and this was not significant (ANOVA, F = 1.06, p = 0.13). We therefore cannot attribute any of the observed variance in allele frequencies directly to environmental variation alone.

However, as the spatial distributions of the two genetic clusters are distinct and there are significant differences in the environments inhabited by the two clusters, the variance in allele frequencies explained by spatial distribution may, in part, be a result of different environmental selection pressures between the two clusters.

Genetic diversity and inbreeding

Across all collection localities observed and expected heterozygosity was very similar (H_O = 0.205-0.232; H_E = 0.242-0.271; Figure 4, Table A1). However, the ANOVA analysis revealed that the eastern genetic cluster had significantly higher observed and expected heterozygosity than the western (p = 0.026 and p=<0.001 respectively Table A5) and admixed collection sites (p = 0.026 and p < 0.081 respectively; Table A5). There was no significant difference in H_O, H_E, and F for the western cluster compared to the admixed collection localities. There were no significant differences found in the inbreeding coefficient (F) among groups (F = 0.107-0.203; Figure 4, Table A1).

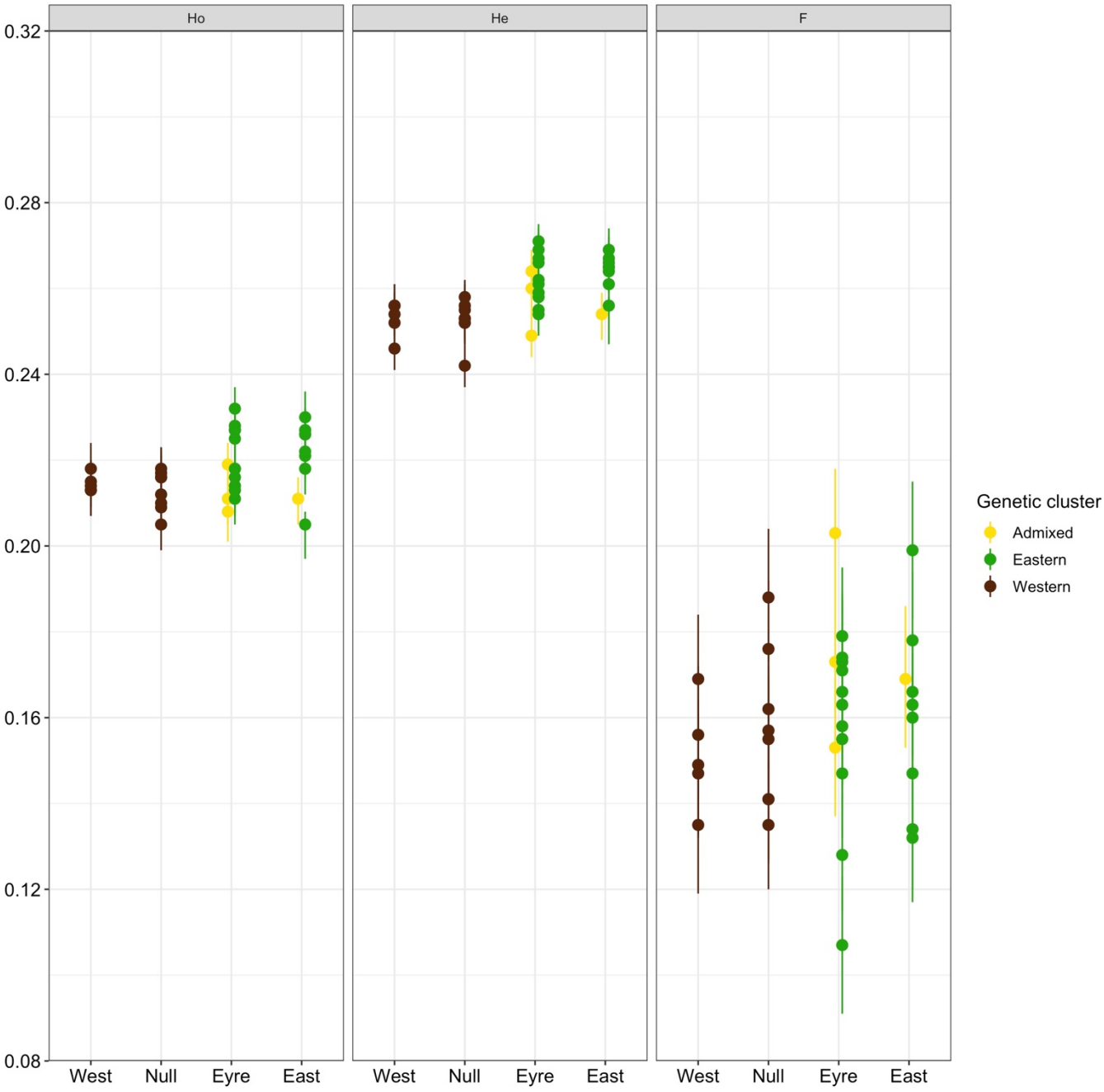


Figure 4. Plots of observed heterozygosity (H_O), expected heterozygosity (H_E) (both vary from 0 to 1, with 0 indicating no heterozygous sites), and inbreeding coefficient (F) (varies from 0 to 1 and values of 0 indicate no evidence of inbreeding). Bars represent 95% confidence intervals. Colours represent genetic cluster assignment identified in the ADMIXTURE analysis and collection localities with >20% admixture (brown: western genetic cluster, yellow: collection localities with >20% admixture, green: eastern genetic cluster). For results of the Tukey post-hoc pairwise comparison refer to Table S2.

Discussion

Overview

We used a genome-wide SNP dataset to conduct a range-wide population genomic study of *Maireana sedifolia,* a keystone restoration species of arid Australia. Overall, the level of population genetic structure was low across the >1500 km range of *M. sedifolia*. Our analysis defined a clear east-west genetic divide, with detectable levels of admixture in all collection localities in the eastern cluster and most pronounced in the most northern samples. Means of genetic diversity significantly differed between the two genetic clusters, which inhabit distinct environments. While our analysis was unable to link any changes in the environment to the genetic differentiation we found, we use this differentiation to define precautionary seed zones. Our findings have implications for seed sourcing for restoration in this species: we suggest that seed can be collected from a far larger area than prescribed under a 'local is best' policy, but the significance of the east-west genetic divide to local adaptation and population fitness requires further exploration.

Genetic structure, diversity, and environmental differences between clusters

The weak levels of population genetic structure observed for *M. sedifolia* are consistent with predictions based on it being a long-lived, wind-pollinated shrub with a large range without any major disjunctions (Hamrick and Godt, 1996, Loveless and Hamrick, 1984). Wind-pollinated species experience gene flow across greater distances than species pollinated by invertebrates, and long-lived species have more mating opportunities and decreased likelihood of genetic drift in populations compared to short-lived species, reducing genetic divergence between populations overall (Broadhurst et al., 2017, Hamrick and Godt, 1996, Loveless and Hamrick, 1984). The weak genetic structure observed is also consistent with the relatively low variation in observed and expected heterozygosity between collection localities within

each genetic cluster, although the eastern genetic cluster had significantly higher genetic diversity overall than the western cluster.

We have achieved some of the most extensive population genomic sampling of a plant species across the Flinders Ranges, Eyre Peninsula, the Gawler Ranges and the Nullarbor Plain, therefore there are limited comparisons we can make. Several published studies of South Australian plant species exist, but they cover a much smaller range than this study. A barrier to gene flow between the Mt Lofty and Flinders Ranges was suggested by studies of *Callistemon teretifolious*, a narrow-range endemic bottlebrush (McCallum et al., 2014), the native short-lived shrub *Goodenia amplexans* Kireta et al. (2019) and *Dodonea viscosa* (Christmas et al., 2017), but we did not find this to be the case for *M. sedifolia.* Interestingly, similarly to *M. sedifolia*, seed and pollen in *D. viscosa* are both considered to be wind dispersed (Harrington and Gadek, 2009), and *D. viscosa* is also widespread with a near continuous range, but there was greater genetic structure found in *D. viscosa* (Christmas et al., 2017). The low structure in *M. sedfiolia* could potentially be due to its long lifespan (Hamrick and Godt, 1996). These previous studies also identified the southern Flinders Ranges as a potential climate refugium, due to elevated levels of genetic diversity in populations found in this region. While for *M. sedifolia*, some of the highly admixed collection localities in the eastern cluster aligned with this region of refugia, we found these sites to have significantly lower levels of genetic diversity than other collection localities. These results highlight the importance of sequencing species in conservation and restoration programmes – for *M. sedifolia,* previously identified barriers to gene-flow (e.g., the Flinders Ranges) do not seem to have caused divergence between populations here and developing seed sourcing guidelines based around previous findings would not have been appropriate.

Whilst we identify clear genetic structure between eastern and western populations, it is not clear what may have driven this divergence. Further investigation of the historical dispersal patterns, genetic drift and selection are

required to more thoroughly understand the low-level divergence we observe across this species' range. Historical dune system movements over the past 0.8 Myr were modelled by Byrne et al. (2008). Caused by huge oscillations in climate (e.g., rainfall, temperature), these dune systems proved inhospitable to plants and caused range disjunctions. One such dune system was located where we find the main location of genetic divergence, near Ceduna. This could plausibly have caused a temporary range disjunction at the location of this genetic break, and the retraction of these dune systems may have led to renewed connectivity between the two clusters. In order to determine this, we would recommend further research in this space through the modelling of the demographic history of the species.

Contemporary environmental factors could also be contributing to the genetic differentiation that we found. While we were unable to show that environmental variables explained any of the observed genetic variation beyond that explained by the spatial distribution of the populations, our analysis did find significant differences between environments of the eastern and western genetic clusters, and between those clusters and localities with >20% admixture. Silt and minimum temperature appeared to be the main variables differentiating the environments of collection localities between the eastern and western genetic clusters. The admixed populations were significantly different from the eastern genetic cluster, and this differentiation was driven by a mixture of soil density, temperature range, aridity, and soil pH, silt content, and aridity index. The genetic break we identify also aligns loosely with the eastern 'Nullarbor Plain Barrier', which is caused predominantly by the high levels of calcrete on the Nullarbor Plain and renders this area inhospitable to plant species with low calcrete tolerance (Crisp and Cook, 2007). Interestingly, the range of *M. sedifolia* spans the Nullarbor and much further east to less calcrete soils. However, our collection localities with high assignment to the western cluster were found exclusively west of this barrier. While there is a general lack of research into the genetic diversity and structure of Australian arid zone plant taxa, Crisp and Cook

(2007) studied causes of high divergence between plant taxa in the southwest and southeast of Australia and implicated the emergence of the highly calcrete Nullarbor Plain 13–14 million years ago. They also found a lack of calcrete tolerance in plant species to increase the likelihood of the Nullarbor creating a barrier to gene flow and thus species divergence. Nullarbor populations of *M. sedifolia* are highly tolerant of calcrete soil and this may be partly driving the genetic structure we have uncovered. We show that the soil conditions (e.g., aridity, pH, silt, organic carbon) are driving the significant differences in environment that we found. Unfortunately, with this set of environmental variables, it is hard to accurately estimate the levels of calcrete found—thus, the question of whether the unique edaphic conditions of the Nullarbor and variable calcrete tolerance of *M. sedifolia* populations have influenced genetic divergence across this barrier requires further research.

Restoration seed zones – management implications

Understanding genetic structure and diversity is vital for the successful delineation of seed zones because it can help to identify areas in the landscape where gene flow has been historically limited, and which populations are inbred or have high genetic diversity (Breed et al., 2018, Breed et al., 2019, Williams et al., 2014). Furthermore, it allows evidence-based seed zones to be defined, which are often more geographically inclusive than seed zones based on field observations (Jørgensen et al., 2016) or arbitrary local boundaries alone. Hence, using population genetic structure information to define seed zones is becoming an accepted way to increase the success of restoration projects by increasing the seed source genetic diversity without increasing the risks of introducing maladapted plants and contaminating local gene pools (Breed et al., 2013, Jørgensen et al., 2016).

Australia has a growing restoration genomics movement, but most keystone restoration species have not had yet been sequenced. In South Australia, for example, a number of genetic studies of plant species exist, but they cover a relatively small

portion of the state and many arid-zone species are excluded (Blyth et al., 2021, Blyth et al., 2020, Christmas et al., 2017, Kireta et al., 2019, McCallum et al., 2013), highlighting the need for further research. There are suggestions in the literature that it is possible to predict genetic diversity and structure based on life-history traits such as reproductive strategy and growth form (Broadhurst et al., 2017, Bussell et al., 2006, Hamrick and Godt, 1996, Jørgensen et al., 2016), and ecological factors such as geographic range (Broadhurst et al., 2017, Loveless and Hamrick, 1984). However, our results suggest that these predictors alone are not enough, and that contemporary and historical environmental variation must be considered. Previous studies demonstrate that barriers to gene flow can be unpredictable (Blyth et al., 2021, Blyth et al., 2020) and that using the genetic structure of a similar species may not be a good proxy (Rossetto et al., 2020).

Perhaps even more importantly, there are fewer studies which identify selection in the genome of plant species and incorporate these findings into seed-sourcing recommendations (but see Prober et al., (2016). Here, we attempted to reveal whether any genetic variation was under divergent environmental selection, but our attempts were confounded by a strong correlation between genetic structure and environmental variation, as well as the lack of a reference genome that would allow us to map the genetic variation. Incorporating evidence of selection and local adaptation into seed sourcing recommendations should therefore be the next step in advancing this field of research (Breed et al., 2019), but reduced-representation genomic data, particularly in the absence of a reference genome, is potentially missing important regions of the genome under selection (Lowry et al., 2017). Therefore, to more fully reveal the genomic landscape of divergence among populations research in this space should utilise whole-genome sequencing combined with mapping to reference, annotated genome. This type of design is becoming an ever more affordable approach, even for restoration projects with small budgets.

Given that we did not identify any populations with relatively low levels of genetic diversity or high inbreeding, the ongoing genetic-diversity management of the species is not of high concern. For now, we recommend the two genetic clusters be employed as preliminary seed zones for *M. sedifolia*, and that seeds should not be transferred across the east-west divide of these clusters. Transferring seed across this genetic break potentially introduces risks such as the introduction of maladapted genes to the local gene pool (such as plants poorly adapted to calcrete soils of the Nullarbor), which could result in poor restoration success. Furthermore, since the admixed collection localities have significant environmental differences to both the eastern and western genetic clusters, we would not recommend using the admixed localities as a seed source for either cluster. For the admixed localities, we recommend only using similarly admixed seed sources until more is understood until more is understood about any functional and adaptive consequences of the admixture we observe in these populations.

Future studies

While we identify clear population genetic structure of *M. sedifolia*, there is considerable scope to expand our understanding of this species and of the population genetic structure of flora in the Australian arid zone in general. We would recommend studies which include taxa which also span this genetic break to determine whether this is a universal break. If so, this could be incorporated into more general seed sourcing guidelines for the region. Although our data can provide insights into genetic diversity and structure, they do not reveal genetic divergence attributable to selection and local adaptation. The semi-random sampling of the genome achieved with reduced representation genomic libraries means that we may be missing a representation of key genomic regions associated with the divergence we observe (Lowry et al., 2017). As such, we would recommend assembling a reference genome and conducting a whole-genome analysis that includes individuals from both genetic clusters and from admixed collection localities. While whole-genome studies are not

currently widely used in restoration studies for financial reasons, infrastructure and resources such as the 10,000 Plant Genomes are in development which will make this a more realistic goal in the near future (Twyford, 2018).

Further knowledge of the implications of mixing seed between the two clusters would also be desirable. Collecting seed from across the species range and growing them in common garden trials would help establish whether there are measurable traits associated with the genetic structure. For example, fine-scale patterns of adaptation (Richardson et al., 2014), could be explored by establishing common garden trials of several populations that span replicated aridity gradients, which would help understand the local functional and potentially adaptive patterns that could have been missed in this study. In addition, this approach would help identify any differences in functional responses (e.g., to drought tolerance), between the clusters, potentially providing the requirements for developing a range-wide climate-adjusted provenancing strategy for restoration (Prober et al., 2015).

Acknowledgements

The authors wish to thank the fantastic team at Jacinth Ambrosia — Jade Zander, Joanne Lee, Tina Law, Brett Dalzell, who all helped coordinate many complex aspects of this remote work and provided expert advice and good company. Thanks go to the horticultural team at the Australian Arid Lands Botanic Gardens, for their expert advice and assistance in our (failed) attempts to grow *Maireana sedifolia*. Thanks go to the enthusiastic team of volunteers who assisted with many remote field trips – Kiri Marker, Jacob Mills, Romy Dresken, Elleca Polson, Sanjaya Rai, Dona Kireta, and to Harrison Sydnor for help in the lab. Thanks also go to the Holsworth Wildlife Research Endowment and the Nature Foundation of South Australia (both awarded to CB), the Australian Research Council (DE150100542 awarded to MFB; DP150103414 awarded to MFB, AJL), and Iluka Resources Limited for funding support.

References

Amor, M.D., Walsh, N.G. and James, E.A., 2020. Genetic Patterns and Climate Modelling Reveal Challenges for Conserving Sclerolaena napiformis (Amaranthaceae sl) an Endemic Chenopod of Southeast Australia. Diversity, 12(11), p.417.

Atlas of Living Australia [Online]. http://www.ala.org.au. Accessed 10 February 2017. [Accessed].

Alexander, D. H., Novembre, J. & Lange, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome research,* 19**,** 1655-1664.

Alpers, D. L., Walker, F. M., Taylor, A. C., Sunnucks, P., Bellman, S., Hansen, B. D. & Sherwin, W. B. 2016. Evidence of subdivisions on evolutionary timescales in a large, declining marsupial distributed across a phylogeographic barrier. *PloS one,* 11**,** e0162789.

Bijlsma, R. and Loeschcke, V., 2012. Genetic erosion impedes adaptive responses to stressful environments. *Evolutionary Applications*, *5*(2), pp.117-129.

Blackwell, W. H. & Powell, M. J. 1981. A preliminary note on pollination in the Chenopodiaceae. *Annals of the Missouri Botanical Garden,* 68**,** 524-526.

Blyth, C., Christmas, M. J., Bickerton, D. C., Breed, M. F., Foster, N. R., Guerin, G. R., Mason, A. R. & Lowe, A. J. 2021. Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy. *Life,* 11**,** 553.

Blyth, C., Christmas, M. J., Bickerton, D. C., Faast, R., Packer, J. G., Lowe, A. J. & Breed, M. F. 2020. Increased Genetic Diversity via Gene Flow Provides Hope for Acacia whibleyana, an Endangered Wattle Facing Extinction. *Diversity,* 12**,** 299.

Breed, M. F. 2013. *Restoration Genetics of Murray Mallee and Neotropical Forests.* PhD thesis. University of Adelaide, Australia.

Breed, M. F., Harrison, P. A., Bischoff, A., Durruty, P., Gellie, N. J., Gonzales, E. K., Havens, K., Karmann, M., Kilkenny, F. F. & Krauss, S. L. 2018. Priority actions to improve provenance decision-making. *BioScience,* 68**,** 510-516.

Breed, M. F., Harrison, P. A., Blyth, C., Byrne, M., Gaget, V., Gellie, N. J., Groom, S. V., Hodgson, R., Mills, J. G. & Prowse, T. A. 2019. The potential of genomics for restoring ecosystems and biodiversity. *Nature Reviews Genetics,* 20**,** 615-628.

Breed, M. F., Mortimer, P. E. & Lowe, A. J. 2016. Restoration: 'Garden of Eden' unrealistic. *Nature,* 533**,** 469.

Breed, M. F., Stead, M. G., Ottewell, K. M., Gardner, M. G. & Lowe, A. J. 2013. Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. *Conservation Genetics,* 14**,** 1-10.

Broadhurst, L., Breed, M. F., Lowe, A., Bragg, J., Catullo, R., Coates, D., Encinas-Viso, F., Gellie, N., James, E., Krauss, S. & Byrne, M. 2017. Genetic diversity and structure of the Australian flora. *Diversity and Distributions,* 23**,** 41-52.

Broadhurst, L., Bush, D. & Begley, J. 2021. Managing Genetic Diversity and Representation in Banksia marginata (Proteaceae) Seed Production Areas Used for Conservation and Restoration. *Diversity,* 13**,** 39.

Broadhurst, L. M., Lowe, A., Coates, D. J., Cunningham, S. A., Mcdonald, M., Vesk, P. A. & Yates, C. 2008. Seed supply for broadscale restoration: maximizing evolutionary potential. *Evolutionary Applications,* 1**,** 587-597.

Bussell, J. D., Hood, P., Alacs, E. A., Dixon, K. W., Hobbs, R. J. & Krauss, S. L. 2006. Rapid genetic delineation of local provenance seed-collection zones for effective rehabilitation of an urban bushland remnant. *Austral Ecology,* 31**,** 164-175.

Butler P.R. 1985. Regeneration studies of M. sedifolia (pearl bluebush) at Manna Hill, South Australia, Natural Resources Project IIA, Graduate Diploma in Natural Resources, Roseworthy Agricultural College (South Australia)

Byrne, M., Yeates, D. K., Joseph, L., Kearney, M., Bowler, J., Williams, M. a. J., Cooper, S., Donnellan, S. C., Keogh, J. S., Leys, R., Melville, J., Murphy, D. J., Porch, N. & Wyrwoll, K. H. 2008. Birth of a

biome: insights into the assembly and maintenance of the Australian arid zone biota. *Molecular Ecology,* 17**,** 4398-4417.

Carvalho, Y. G., Vitorino, L. C., De Souza, U. J. & Bessa, L. A. 2019. Recent trends in research on the genetic diversity of plants: implications for conservation. *Diversity,* 11**,** 62.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. H. 2011. Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics,* 1**,** 171-182.

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. 2013. Stacks: an analysis tool set for population genomics. *Molecular ecology,* 22**,** 3124-3140.

Christmas, M., Biffin, E., Breed, M. & Lowe, A. 2017. Targeted capture to assess neutral genomic variation in the narrow-leaf hopbush across a continental biodiversity refugium. *Scientific Reports,* 7**,** Publication number 41367.

Christmas, M.J., Breed, M.F. and Lowe, A.J., 2016. Constraints to and conservation implications for climate change adaptation in plants. *Conservation Genetics*, *17*(2), pp.305-320.

Clewell, A. & Rieger, J. P. 1997. What practitioners need from restoration ecologists. *Restoration Ecology,* 5, 350-354.

Crisp, M.D. 1978. Demography and survival under grazing of three Australian semi-desert shrubs, *Oikos* 30 520-528

Crisp, M. D. & Cook, L. G. 2007. A congruent molecular signature of vicariance across multiple plant lineages. *Molecular phylogenetics and evolution,* 43**,** 1106-1117.

Crouzeilles, R., Curran, M., Ferreira, M.S., Lindenmayer, D.B., Grelle, C.E. and Benayas, J.M.R., 2016. A global meta-analysis on the ecological drivers of forest restoration success. Nature communications, 7(1), pp.1-8.

Cunningham G.M., Mulham W.E., Milthorpe P.L., Leigh J.H. 1981. *Plants of western New South Wales* Soil. Conservation Service of New South Wales, Government Printer, Sydney.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., Depristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T. & Sherry, S. T. 2011. The variant call format and VCFtools. *Bioinformatics,* 27**,** 2156-2158.

De Kort, H., Mergeay, J., Vander Mijnsbrugge, K., Decocq, G., Maccherini, S., Kehlet Bruun, H. H., Honnay, O. & Vandepitte, K. 2014. An evaluation of seed zone delineation using phenotypic and population genomic data on black alder Alnus glutinosa. *Journal of applied ecology,* 51**,** 1218-1227.

de Caritat, P., Cooper, M. and Wilford, J., 2011. The pH of Australian soils: field results from a national survey. *Soil Research*, *49*(2), pp.173-182.

Dixon, P. 2003. VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science,* 14**,** 927-930.

Dolman, G. & Joseph, L. 2015. Evolutionary history of birds across southern Australia: structure, history and taxonomic implications of mitochondrial DNA diversity in an ecologically diverse suite of species. *Emu,* 115**,** 35-48.

Feng, X., Fu, B., Lu, N., Zeng, Y. & Wu, B. 2013. How ecological restoration alters ecosystem services: an analysis of carbon sequestration in China's Loess Plateau. *Scientific reports,* 3**,** 2846.

Frankham, R. 2005. Genetics and extinction. *Biological Conservation,* 126**,** 131-140.

Gallant, J., Austin, J., Williams, K., Harwood, T., King, D., Nolan, M., Mokany, K. & Csiro 2018. 9s soil and landform for continental Australia analysis of biodiversity pattern: aggregated from 3s data. v1. *Data collection*.

Gellie, N. J., Breed, M. F., Thurgate, N., Kennedy, S. A., & Lowe, A. J. (2016). Local maladaptation in a foundation tree species: Implications for restoration. Biological conservation, 203, 226-232.

Gibbs, H. K. & Salmon, J. M. 2015. Mapping the world's degraded lands. *Applied Geography,* 57**,** 12-21.

Godefroid, S., Piazza, C., Rossi, G., Buord, S., Stevens, A.-D., Aguraiuja, R., Cowell, C., Weekley, C. W., Vogg, G., Iriondo, J. M., Johnson, I., Dixon, B., Gordon, D., Magnanon, S., Valentin, B., Bjureke, K., Koopman, R., Vicens, M., Virevaire, M. & Vanderborght, T. 2011. How successful are plant species reintroductions? *Biological Conservation,* 144**,** 672-682.

Grundy, M., Rossel, R. V., Searle, R., Wilson, P., Chen, C. & Gregory, L. 2015. Soil and landscape grid of Australia. *Soil Research,* 53**,** 835-844.

Hamrick, J. L. & Godt, M. J. W. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society B-Biological Sciences,* 351**,** 1291-1298.

Harrington, M. G. & Gadek, P. A. 2009. A species well travelled – the *Dodonaea viscosa* (Sapindaceae) complex based on phylogenetic analyses of nuclear ribosomal ITS and ETSf sequences. *Journal of Biogeography,* 36**,** 2313-2323.

Harwood, T., Donohue, R., Harman, I., Mcvicar, T., Ota, N., Perry, J., Williams, K. & CSIRO. 2016. 9s climatology for continental Australia 1976-2005: Summary variables with elevation and radiative adjustment. v1. *Data collection*.

Hijmans, R. J. 2021. Geographic Data Analysis and Modeling [R package raster version 3.4-10].

Hufford, K. M. & Mazer, S. J. 2003. Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution,* 18**,** 147-155.

Huson, D. H. 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics (Oxford, England),* 14**,** 68-73.

Jombart, T., Devillard, S. & Balloux, F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics,* 11**,** 94.

Jørgensen, M. H., Elameen, A., Hofman, N., Klemsdal, S., Malaval, S. & Fjellheim, S. 2016. What's the meaning of local? Using molecular markers to define seed transfer zones for ecological restoration in Norway. *Evolutionary applications,* 9**,** 673-684.

Kireta, D., Christmas, M. J., Lowe, A. J. & Breed, M. F. 2019. Disentangling the evolutionary history of three related shrub species using genome-wide molecular markers. *Conservation Genetics,* 20**,** 1101-1112.

Krauss, S. L., Sinclair, E. A., Bussell, J. D. & Hobbs, R. J. 2013. An ecological genetic delineation of local seed-source provenance for ecological restoration. *Ecology and evolution,* 3**,** 2138-2149.

Loveless, M. D. & Hamrick, J. L. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics,* 15**,** 65-95.

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F. & Storfer, A. 2017. Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Wiley Online Library.

McCallum, K.P., Guerin, G.R., Breed, M.F. and Lowe, A.J., 2014. Combining population genetics, species distribution modelling and field assessments to understand a species vulnerability to climate change. *Austral Ecology*, *39*(1), pp.17-28.

McDonald J, Breed M.F., Gellie N.G., Blyth C, Lowe AJ. 2017. A life cycle review of Maireana sedifolia with a focus on seeding triggers. Technical Report prepared by The University of Adelaide, on behalf of Iluka Resources.

Mckay, J. K., Christian, C. E., Harrison, S. & Rice, K. J. 2005. "How local is local?" - a review of practical and conceptual issues in the genetics of restoration. *Restoration Ecology,* 13**,** 432-440.

Meirmans, P. G. 2020. genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular ecology resources,* 20**,** 1126-1131.

Meirmans, P. G. & Van Tienderen, P. H. 2004. GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes,* 4**,** 792-794.

Neaves, L. E., Zenger, K. R., Prince, R. I., Eldridge, M. D. & Cooper, D. W. 2009. Landscape discontinuities influence gene flow and genetic structure in a large, vagile Australian mammal, Macropus fuliginosus. *Molecular Ecology,* 18**,** 3363-3378.

Nkonya, E., Mirzabaev, A. & Von Braun, J. 2016. Economics of land degradation and improvement: An introduction and overview. *Economics of Land Degradation and Improvement – A Global Assessment for Sustainable Development.* Springer, Cham.

Pembleton, L. W., Cogan, N. O. & Forster, J. W. 2013. St AMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular ecology resources,* 13**,** 946-952.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE,* 7**,** e37135.

Pike, C., Potter, K.M., Berrang, P., Crane, B., Baggs, J., Leites, L. and Luther, T., 2020. New seed-collection zones for the eastern United States: the eastern seed zone forum. *Journal of Forestry*, *118*(4), pp.444-451

Prober, S. M., Byrne, M., Mclean, E. H., Steane, D. A., Potts, B. M., Vaillancourt, R. E. & Stock, W. D. 2015. Climate-adjusted provenancing: a strategy for climate-resilient ecological restoration. *Frontiers in Ecology and Evolution,* 3, p65.

Prober, S. M., Potts, B. M., Bailey, T., Byrne, M., Dillon, S., Harrison, P. A., Hoffmann, A. A., Jordan, R., Mclean, E. H. & Steane, D. A. 2016. Climate adaptation and ecological restoration in eucalypts. *Proceedings of the Royal Society of Victoria,* 128**,** 40-53.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. & Daly, M. J. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics,* 81**,** 559-575.

Richardson, J. L., Urban, M. C., Bolnick, D. I. & Skelly, D. K. 2014. Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution,* 29**,** 165-176.

Rosenberg, N. A. 2004. DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes,* 4**,** 137-138.

Rossetto, M., Wilson, P. D., Bragg, J., Cohen, J., Fahey, M., Yap, J.-Y. S. & Van Der Merwe, M. 2020. Perceptions of Similarity Can Mislead Provenancing Strategies—An Example from Five Co-Distributed Acacia Species. *Diversity,* 12**,** 306.

Ser 2004. The SER international primer on ecological restoration, Version 2. Tucson, Arizona.: Society for Ecological Restoration Science and Policy Working Group.

Shafer, A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W. & Wolf, J. B. 2016. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*.

Specht R.L (1972) The vegetation of South Australia, Australian Government Press, Adelaide

R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing.

Van Woerdon, S. 2014. Declaration would end billions of tons of climate pollution per year, restore 350 million hectares of forest; backed by tangible private sector commitments. New York: United Nations.

Twyford, A.D., 2018. The road to 10,000 plant genomes. *Nature plants*, *4*(6), pp.312-313.

Williams, A. V., Nevill, P. G. & Krauss, S. L. 2014. Next generation restoration genetics: applications and opportunities. *Trends in Plant Science,* 19**,** 529-537.

Appendix

Table A1. Summary of region, collection locality, genetic cluster (localities with >20% admixture between the two genetic clusters are identified), sample size after filtering (n), observed (H_O) and expected (H_E) heterozygosity (both vary from 0 to 1, with 0 indicating no diversity), inbreeding coefficient (F) (varies from 0 to 1 and values of 0 indicate no evidence of inbreeding.

Region	Collection locality	Genetic cluster	Latitude	Longitude	n	*Ho*	H_E	*F*
Western Australia	West-A	Western	-29.66926	125.13025	8	0.213	0.256	0.169
Western Australia	West-B	Western	-31.01671	125.32921	8	0.215	0.252	0.147
Western Australia	West-C	Western	-29.47343	127.07193	7	0.218	0.256	0.149
Western Australia	West-D	Western	-32.01217	126.26339	7	0.214	0.254	0.156
Western Australia	West-E	Western	-31.95245	127.27428	7	0.213	0.246	0.135
Nullarbor	Null-A	Western	-29.39064	130.53213	8	0.217	0.253	0.141
Nullarbor	Null-B	Western	-30.51416	131.81604	7	0.216	0.256	0.157
Nullarbor	Null-C	Western	-31.07743	131.77766	7	0.205	0.252	0.188
Nullarbor	Null-D	Western	-30.48098	132.10196	8	0.212	0.253	0.162
Nullarbor	Null-E	Western	-30.8859	132.19722	8	0.218	0.258	0.155
Nullarbor	Null-F	Western	-31.63402	132.14169	8	0.216	0.252	0.141
Nullarbor	Null-G	Western	-31.24243	132.58429	6	0.21	0.255	0.176
Nullarbor	Null-H	Western	-31.86687	132.86456	8	0.209	0.242	0.135
Eyre Peninsula/Gawler Ranges	Eyre-A	Admixed	-30.86315	135.05634	8	0.219	0.264	0.173
Eyre Peninsula/Gawler Ranges	Eyre-B	Admixed	-31.25796	135.25032	8	0.208	0.26	0.203
Eyre Peninsula/Gawler Ranges	Eyre-C	Eastern	-31.93626	135.20193	8	0.216	0.262	0.174
Eyre Peninsula/Gawler Ranges	Eyre-D	Eastern	-32.60603	135.27699	7	0.214	0.259	0.171
Eyre Peninsula/Gawler Ranges	Eyre-E	Eastern	-32.52636	136.50493	8	0.218	0.258	0.155
Eyre Peninsula/Gawler Ranges	Eyre-F	Eastern	-32.68022	135.88843	7	0.218	0.261	0.166
Eyre Peninsula/Gawler Ranges	Eyre-G	Eastern	-33.02094	136.29440	7	0.227	0.254	0.107
Eyre Peninsula/Gawler Ranges	Eyre-H	Admixed	-31.24729	136.56369	7	0.211	0.249	0.153
Eyre Peninsula/Gawler Ranges	Eyre-I	Eastern	-32.71974	137.14795	8	0.232	0.266	0.128
Eyre Peninsula/Gawler Ranges	Eyre-J	Eastern	-33.70846	136.82596	8	0.211	0.255	0.173
Eyre Peninsula/Gawler Ranges	Eyre-K	Eastern	-33.34316	137.08932	7	0.213	0.259	0.179
Eyre Peninsula/Gawler Ranges	Eyre-L	Eastern	-32.2403	137.54623	8	0.225	0.269	0.163
Eyre Peninsula/Gawler Ranges	Eyre-M	Eastern	-32.63864	137.54179	8	0.228	0.271	0.158
Eyre Peninsula/Gawler Ranges	Eyre-N	Eastern	-33.01543	137.55446	7	0.227	0.267	0.147
Flinders Ranges	East-A	Eastern	-32.60308	137.89864	8	0.227	0.261	0.132
Flinders Ranges	East-B	Eastern	-32.99147	138.04457	7	0.205	0.256	0.199
Flinders Ranges	East-C	Eastern	-32.6461	138.54179	7	0.218	0.265	0.178
Flinders Ranges	East-D	Admixed	-30.46971	139.30918	6	0.211	0.254	0.169
Flinders Ranges	East-E	Eastern	-33.66895	139.05966	7	0.226	0.269	0.16
Flinders Ranges	East-F	Eastern	-32.62964	139.62337	8	0.222	0.267	0.166
Flinders Ranges	East-G	Eastern	-34.07301	139.25456	8	0.23	0.266	0.134
Flinders Ranges	East-H	Eastern	-34.02285	139.76387	8	0.221	0.264	0.163
Flinders Ranges	East-I	Eastern	-32.29058	140.29708	8	0.226	0.265	0.147

Table A2. Output of the ADMIXTURE analysis, specifically to explore the genetic cluster (K) assignment of individuals with >20% admixture between the two genetic clusters. As the confidence intervals of the highly admixed populations did not overlap with the proportion assignment of 1 or 0, we concluded that the populations were indeed intermediate

Individual	Proportion assigned to K1	Proportion assigned to K2	SE	95% CI	K1-95% CI	K1+95% CI	K2-95% CI	K2+95% CI
East-D-1	0.248264	0.751736	0.026932	0.053864	0.194	0.302128	0.697872	0.8056
East-D-3	0.391394	0.608606	0.028489	0.056978	0.334416	0.448372	0.551628	0.665584
East-D-4	0.52475	0.47525	0.021253	0.042506	0.482244	0.567256	0.432744	0.517756
East-D-5	0.270941	0.729059	0.014447	0.028894	0.242047	0.299835	0.700165	0.757953
East-D-7	0.498181	0.501819	0.02198	0.04396	0.454221	0.542141	0.457859	0.545779
East-D-8	0.441348	0.558652	0.03411	0.06822	0.373128	0.509568	0.490432	0.626872
East-G-7	0.231464	0.768536	0.013562	0.027124	0.20434	0.258588	0.741412	0.79566
East-I-7	0.213566	0.786434	0.01032	0.02064	0.192926	0.234206	0.765794	0.807074
Eyre-A-1	0.640613	0.359387	0.018478	0.036956	0.603657	0.677569	0.322431	0.396343
Eyre-A-2	0.680979	0.319021	0.01608	0.03216	0.648819	0.713139	0.286861	0.351181
Eyre-A-3	0.733977	0.266023	0.012411	0.024822	0.709155	0.758799	0.241201	0.290845
Eyre-A-4	0.610508	0.389492	0.030321	0.060642	0.549866	0.67115	0.32885	0.450134
Eyre-A-5	0.55791	0.44209	0.014245	0.02849	0.52942	0.5864	0.4136	0.47058
Eyre-A-6	0.695347	0.304653	0.015897	0.031794	0.663553	0.727141	0.272859	0.336447
Eyre-A-7	0.738541	0.261459	0.01881	0.03762	0.700921	0.776161	0.223839	0.299079
Eyre-A-8	0.659483	0.340517	0.023254	0.046508	0.612975	0.705991	0.294009	0.387025
Eyre-B-1	0.432742	0.567258	0.013418	0.026836	0.405906	0.459578	0.540422	0.594094
Eyre-B-2	0.487383	0.512617	0.026352	0.052704	0.434679	0.540087	0.459913	0.565321
Eyre-B-3	0.567295	0.432705	0.016853	0.033706	0.533589	0.601001	0.398999	0.466411
Eyre-B-4	0.591572	0.408428	0.029085	0.05817	0.533402	0.649742	0.350258	0.466598
Eyre-B-5	0.649807	0.350193	0.012984	0.025968	0.623839	0.675775	0.324225	0.376161
Eyre-B-6	0.627748	0.372252	0.01979	0.03958	0.588168	0.667328	0.332672	0.411832
Eyre-B-7	0.487928	0.512072	0.015596	0.031192	0.456736	0.51912	0.48088	0.543264
Eyre-B-8	0.516572	0.483428	0.015172	0.030344	0.486228	0.546916	0.453084	0.513772
Eyre-C-1	0.378655	0.621345	0.017166	0.034332	0.344323	0.412987	0.587013	0.655677
Eyre-C-3	0.265102	0.734898	0.023317	0.046634	0.218468	0.311736	0.688264	0.781532
Eyre-C-4	0.271229	0.728771	0.013125	0.02625	0.244979	0.297479	0.702521	0.755021
Eyre-C-5	0.316463	0.683537	0.041254	0.082508	0.233955	0.398971	0.601029	0.766045
Eyre-C-6	0.299725	0.700275	0.030808	0.061616	0.238109	0.361341	0.638659	0.761891
Eyre-C-7	0.223707	0.776293	0.023893	0.047786	0.175921	0.271493	0.728507	0.824079
Eyre-C-8	0.252843	0.747157	0.036409	0.072818	0.180025	0.325661	0.674339	0.819975
Eyre-E-6	0.270628	0.729372	0.044435	0.08887	0.181758	0.359498	0.640502	0.818242
Eyre-F-3	0.253508	0.746492	0.02204	0.04408	0.209428	0.297588	0.702412	0.790572
Eyre-F-6	0.245886	0.754114	0.009567	0.019134	0.226752	0.26502	0.73498	0.773248
Eyre-H-1	0.302696	0.697304	0.021937	0.043874	0.258822	0.34657	0.65343	0.741178
Eyre-H-2	0.487584	0.512416	0.02747	0.05494	0.432644	0.542524	0.457476	0.567356
Eyre-H-3	0.602892	0.397108	0.015969	0.031938	0.570954	0.63483	0.36517	0.429046
Eyre-H-4	0.646917	0.353083	0.025777	0.051554	0.595363	0.698471	0.301529	0.404637
Eyre-H-5	0.479291	0.520709	0.031519	0.063038	0.416253	0.542329	0.457671	0.583747
Eyre-H-6	0.310289	0.689711	0.029141	0.058282	0.252007	0.368571	0.631429	0.747993
Eyre-H-8	0.483765	0.516235	0.033391	0.066782	0.416983	0.550547	0.449453	0.583017
Eyre-I-8	0.204611	0.795389	0.018288	0.036576	0.168035	0.241187	0.758813	0.831965
Eyre-K-7	0.315784	0.684216	0.014102	0.028204	0.28758	0.343988	0.656012	0.71242
Eyre-L-3	0.206928	0.793072	0.029494	0.058988	0.14794	0.265916	0.734084	0.85206
Eyre-L-8	0.367889	0.632111	0.014339	0.028678	0.339211	0.396567	0.603433	0.660789
Eyre-N-5	0.205829	0.794171	0.014447	0.028894	0.176935	0.234723	0.765277	0.823065

Table A3. Pairwise genetic differentiation (measured by FST) between collection localities (values vary from 0–1 and values of 0 indicate no evidence of differentiation).

	West-A	West-B	West-C	West-D	West-E	Null-A	Null-B	Null-C	Null-D	Null-E	Null-F	Null-G	Null-H	Eyre-A	Eyre-B	Eyre-C	Eyre-D	Eyre-E	Eyre-F	Eyre-G	Eyre-H	Eyre-I	Eyre-J	Eyre-K	Eyre-L	Eyre-M	Eyre-N	East-A	East-B	East-C	East-D	East-E	East-F	East-G	East-H	East-I
West-A	0.00																																		
West-B	0.01	0.00																																	
West-C	0.00	0.01	0.00																																
West-D	0.00	0.01	0.01	0.00																															
West-E	0.01	0.02	0.01	0.01	0.00																														
Null-A	0.01	0.03	0.01	0.02	0.02	0.00																													
Null-B	0.01	0.02	0.01	0.02	0.02	0.02	0.00																												
Null-C	0.00	0.02	0.01	0.01	0.01	0.02	0.00	0.00																											
Null-D	0.01	0.02	0.01	0.01	0.02	0.01	0.01	0.00	0.00																										
Null-E	0.00	0.02	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.00																									
Null-F	0.02	0.03	0.02	0.03	0.03	0.03	0.02	0.02	0.03	0.01	0.00																								
Null-G	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.00	0.01	0.00	0.02	0.00																							
Null-H	0.03	0.05	0.04	0.04	0.04	0.04	0.04	0.03	0.03	0.03	0.04	0.02	0.00																						
Eyre-A	0.01	0.03	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.03	0.01	0.04	0.00																					
Eyre-B	0.02	0.04	0.03	0.03	0.04	0.03	0.03	0.03	0.03	0.02	0.04	0.02	0.05	0.01	0.00																				
Eyre-C	0.03	0.04	0.03	0.03	0.04	0.04	0.03	0.03	0.03	0.03	0.04	0.03	0.05	0.02	0.02	0.00																			
Eyre-D	0.04	0.05	0.04	0.04	0.05	0.05	0.05	0.04	0.05	0.04	0.05	0.04	0.07	0.02	0.03	0.01	0.00																		
Eyre-E	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.06	0.05	0.07	0.03	0.04	0.01	0.01	0.00																	
Eyre-F	0.04	0.05	0.04	0.04	0.05	0.05	0.04	0.04	0.04	0.04	0.05	0.04	0.06	0.02	0.02	0.00	0.01	0.02	0.00																
Eyre-G	0.06	0.07	0.06	0.07	0.06	0.07	0.07	0.06	0.06	0.06	0.07	0.06	0.09	0.04	0.05	0.03	0.04	0.04	0.03	0.00															
Eyre-H	0.05	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.06	0.06	0.08	0.04	0.05	0.04	0.05	0.05	0.04	0.07	0.00														
Eyre-I	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.06	0.02	0.02	0.00	0.01	0.02	0.00	0.03	0.05	0.00													
Eyre-J	0.06	0.07	0.06	0.07	0.07	0.06	0.07	0.06	0.07	0.06	0.07	0.06	0.09	0.04	0.05	0.03	0.03	0.04	0.03	0.05	0.07	0.03	0.00												
Eyre-K	0.04	0.05	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.04	0.05	0.04	0.06	0.02	0.03	0.01	0.01	0.02	0.01	0.03	0.05	0.01	0.03	0.00											
Eyre-L	0.04	0.04	0.03	0.04	0.04	0.04	0.04	0.03	0.04	0.03	0.04	0.03	0.06	0.02	0.02	0.00	0.02	0.01	0.00	0.03	0.04	0.00	0.03	0.01	0.00										
Eyre-M	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.03	0.05	0.04	0.06	0.02	0.03	0.01	0.01	0.01	0.01	0.03	0.04	0.00	0.03	0.01	0.00	0.00									
Eyre-N	0.04	0.05	0.03	0.04	0.05	0.05	0.04	0.04	0.04	0.04	0.05	0.04	0.06	0.03	0.03	0.01	0.02	0.02	0.01	0.03	0.05	0.00	0.03	0.01	0.01	0.00	0.00								
East-A	0.05	0.06	0.05	0.05	0.06	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.07	0.03	0.04	0.02	0.02	0.02	0.02	0.04	0.05	0.02	0.04	0.02	0.01	0.01	0.02	0.00							
East-B	0.05	0.06	0.05	0.06	0.06	0.06	0.06	0.05	0.05	0.05	0.06	0.05	0.07	0.03	0.04	0.02	0.02	0.02	0.02	0.04	0.05	0.02	0.04	0.01	0.01	0.01	0.02	0.02	0.00						
East-C	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.04	0.06	0.02	0.03	0.01	0.01	0.01	0.00	0.03	0.04	0.00	0.03	0.01	0.00	0.00	0.01	0.01	0.01	0.00					
East-D	0.04	0.05	0.05	0.05	0.06	0.05	0.05	0.04	0.05	0.05	0.06	0.05	0.08	0.03	0.04	0.03	0.04	0.05	0.04	0.06	0.06	0.04	0.06	0.05	0.03	0.04	0.04	0.05	0.04	0.03	0.00				
East-E	0.04	0.05	0.04	0.04	0.04	0.05	0.04	0.04	0.04	0.04	0.04	0.04	0.06	0.02	0.02	0.00	0.01	0.02	0.00	0.03	0.04	0.01	0.02	0.01	0.00	0.01	0.01	0.01	0.01	0.00	0.03	0.00			
East-F	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.05	0.07	0.03	0.03	0.01	0.01	0.02	0.01	0.04	0.05	0.01	0.03	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.04	0.00	0.00		
East-G	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.06	0.04	0.07	0.03	0.04	0.02	0.02	0.02	0.01	0.04	0.06	0.01	0.04	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.04	0.01	0.01	0.00	
East-H	0.04	0.05	0.04	0.04	0.05	0.05	0.04	0.04	0.05	0.04	0.05	0.04	0.07	0.02	0.03	0.01	0.01	0.02	0.01	0.03	0.04	0.01	0.03	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.03	0.00	0.00	0.01	0.00
East-I	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.04	0.04	0.04	0.05	0.04	0.07	0.02	0.03	0.02	0.02	0.02	0.01	0.04	0.05	0.02	0.03	0.02	0.01	0.02	0.02	0.02	0.02	0.01	0.04	0.00	0.01	0.01	0.01	0.00

Table A4. Results of the Tukey pairwise comparison to detect differences in PC1 and PC2 between the three genetic groups between the western genetic cluster, the eastern genetic cluster, and collection localities with >20% admixture. Diff is the difference between means of the two groups, lwr and upr are the ends of the 95% confidence interval and p adj is the p-value after adjusting for the comparisons. We have compared the differences between PC1 (aridity, precipitation, temperature range, soil phosphorus, soil nitrogen, soil density), and PC2 (soil carbon, minimum temperature, and soil silt).

	Group cluster comparison	diff	lwr	upr	*p* adj
PC1: aridity, phosphorous, pH, soil density	Eastern-Admixed	3.076	0.712	5.440	0.008
	Western-Admixed	2.149	-0.307	4.606	0.096
	Western-Eastern	-0.927	-2.473	0.620	0.318
PC2: silt, minimum temperature, nitrogen, carbon	Eastern-Admixed	-1.21	-2.80	0.38	0.17
	Western-Admixed	0.88	-0.78	2.53	0.40
	Western-Eastern	2.08	-1.04	3.13	< 0.001

Table A5. Results of the Tukey post-hoc pairwise comparison to detect differences in observed heterozygosity (H_E), expected heterozygosity (H_E) and the inbreeding coefficient (F) between the western genetic cluster, the eastern genetic cluster, and collection localities with >20% admixture. Diff is the difference between means of the two groups, lwr and upr are the ends of the 95% confidence interval and p adj is the p-value after adjusting for the comparisons.

	Group cluster comparison	diff	lwr	upr	*p* adj
H_O	Eastern-Admixed	0.009	0.001	0.017	0.026
	Western-Admixed	0.001	-0.007	0.010	0.925
	Western-Eastern	-0.008	-0.013	-0.002	0.003
H_E	Eastern-Admixed	0.006	-0.001	0.013	0.081
	Western-Admixed	-0.004	-0.011	0.003	0.338
	Western-Eastern	-0.010	-0.015	-0.006	< 0.001
F	Eastern-Admixed	-0.017	-0.043	0.010	0.287
	Western-Admixed	-0.020	-0.047	0.008	0.198
	Western-Eastern	-0.003	-0.021	0.014	0.893

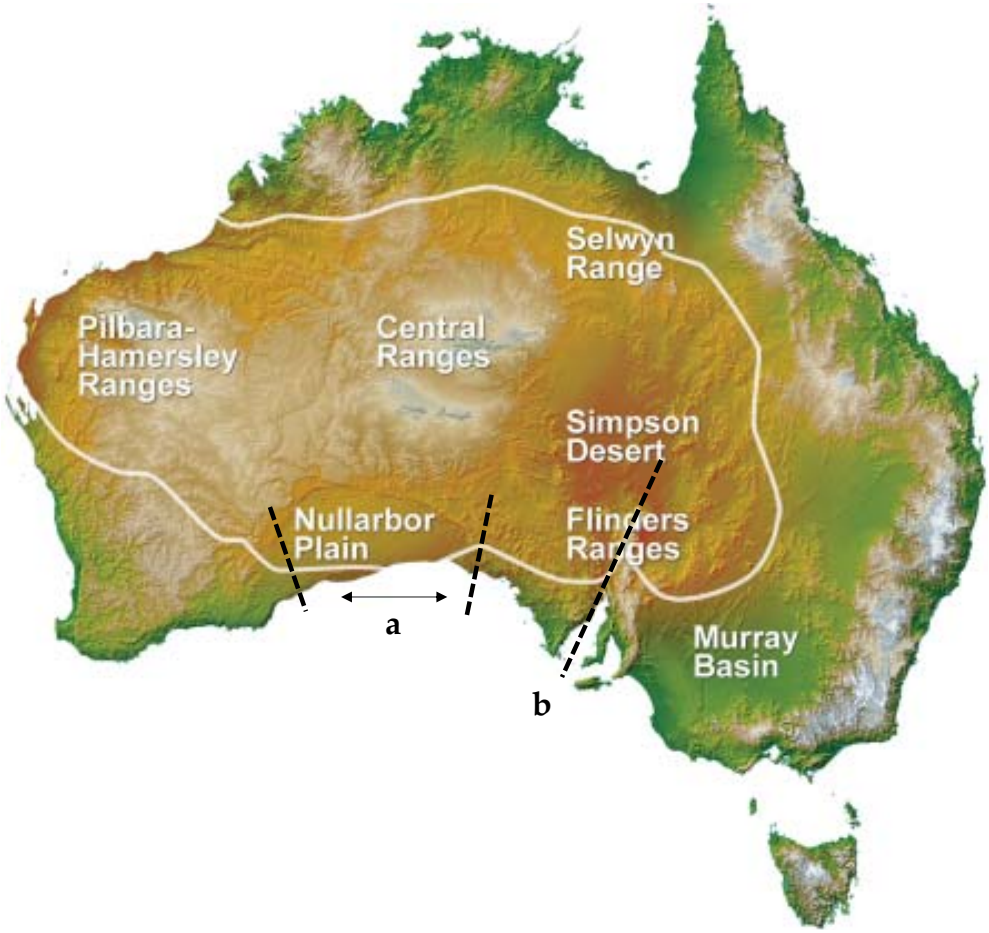
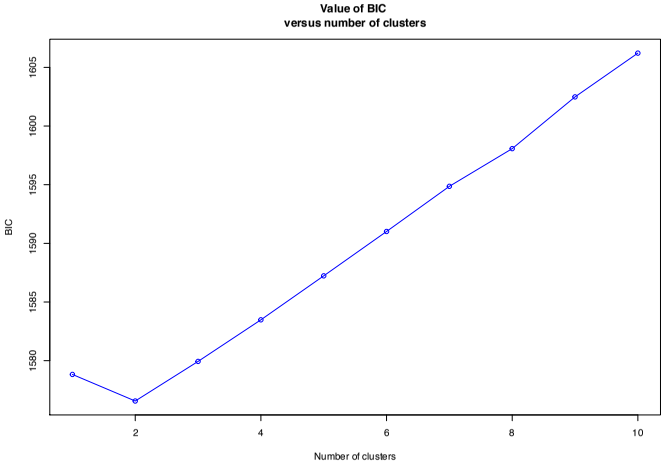


Figure A1. Adapted from Byrne et al. (2008) and Neaves et al. (2009). The map depicts the extent of the arid zone (white line) and the geographical features of the Australian arid zone (e.g., the Flinders Ranges). Major known historical barriers (a. the Nullarbor Plains Barrier and b. the Eyrean Barrier) are identified by the black dotted line.

(a)



(b)

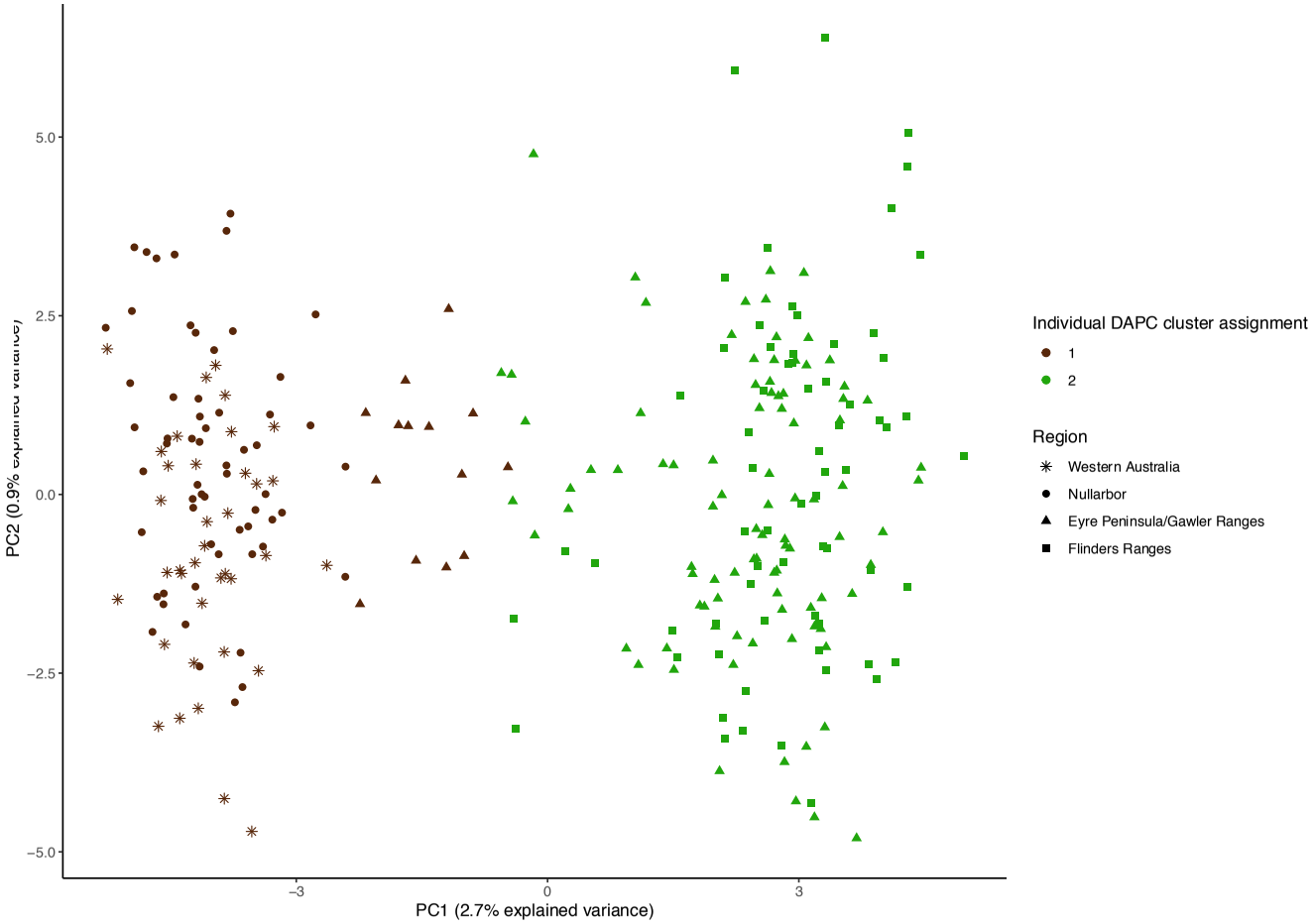


Figure A2. Discriminant analysis of principal components (DAPC) Bayesian Information Criterion results. (a)The optimal number of clusters (K) as determined by 'k-means', a clustering algorithm that looks for the value of *K* that maximises the variation between groups. The Bayesian Information Criterion (BIC) is plotted for K = 1-10, and the 'elbow' in the graph at K = 1-2 indicates these to be the most likely value of K. (b) Principal Component Analysis showing population genetic structure across all samples. Colours represent individual genetic cluster assignment from the DAPC analysis. Shape represents collection region.

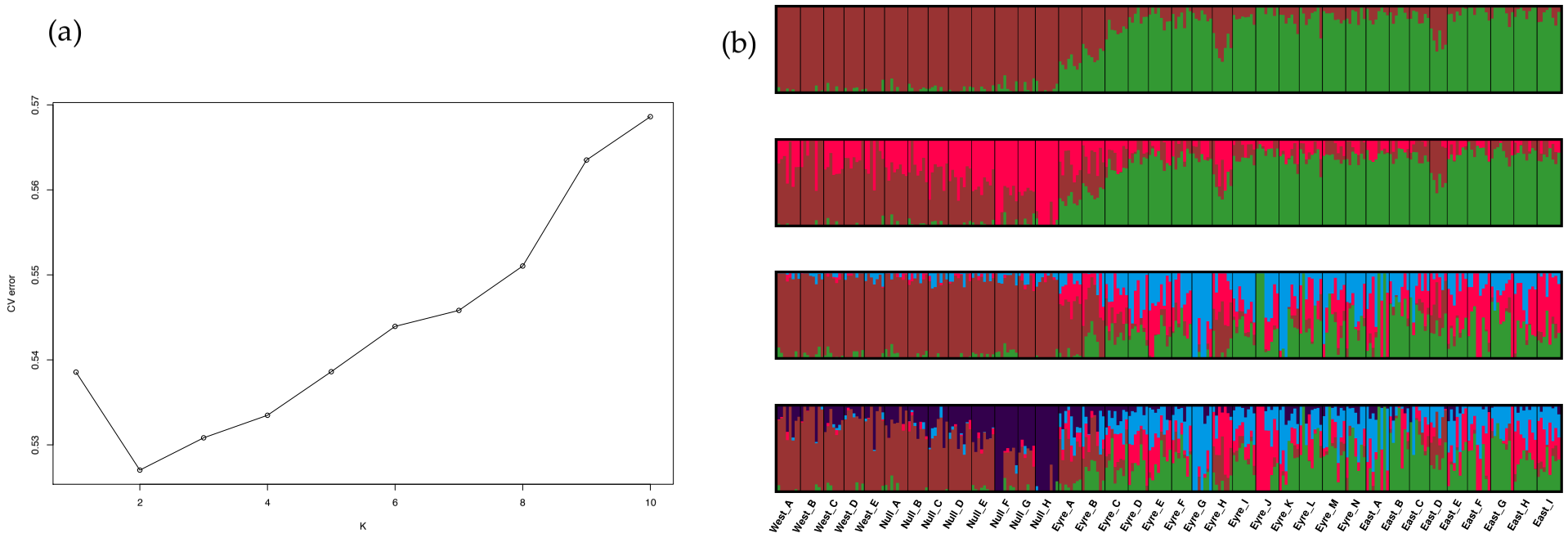


Figure A3. ADMIXTURE error results—the cv errors are plotted for K = 1–10, with the lowest cv error indicating greatest support at K = 6. (b) Bar plots represent individual genetic cluster assignment from ADMIXTURE results from K = 2 to K = 5).

Statement of Authorship

Title of Paper	Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy
Publication Status	■ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Blyth, C., Christmas, M.J., Bickerton, D.C., Breed, M.F., Foster, N.R., Guerin, G.R., Mason, A.R. and Lowe, A.J., 2021. Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy. *Life*, *11*(6), p.553.

Principal Author

Name of Principal Author (Candidate)	Colette Blyth		
Contribution to the Paper	Helped develop the ideas and methods, conducted fieldwork, analysed the data and wrote the manuscript.		
Overall percentage (%)	60		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	8/7/2021

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Dr Matt Christmas		
Contribution to the Paper	Helped develop the ideas and methods, helped analyse the data, contributed to draft revisions of manuscript.		
Signature		Date	6/7/2021
Name of Co-Author	Doug Bickerton		

Contribution to the Paper	Helped develop the ideas and methods, helped with fieldwork, contributed to draft revisions of manuscript.		
Signature		Date	2/7/21
Name of Co-Author	Dr Martin Breed		
Contribution to the Paper	Helped secure funding, helped develop the ideas and methods, helped analyse the data, contributed to draft revisions of manuscript.		
Signature		Date	2/7/2021
Name of Co-Author	Nicole Foster		
Contribution to the Paper	Helped analyse the data and contributed to draft revisions of manuscript.		
Signature		Date	01/07/2021
Name of Co-Author	Dr Greg Guerin		
Contribution to the Paper	Helped analyse the data and contributed to draft revisions of manuscript.		
Signature		Date	2/7/2021
Name of Co-Author	Alexander Mason		
Contribution to the Paper	Helped with fieldwork, carried out leaf trait measurements, contributed to draft revisions of manuscript.		
Signature		Date	5/7/2021
Name of Co-Author	Professor Andy Lowe		
Contribution to the Paper	Helped secure funding, helped develop the ideas and methods, contributed to draft revisions of manuscript.		
Signature		Date	5/7/2021

Chapter 4. Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy

life

MDPI

Article

Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy

Colette Blyth [1,*], Matthew J. Christmas [2,*], Douglas C. Bickerton [3], Martin F. Breed [4], Nicole R. Foster [1,4], Greg R. Guerin [1,5], Alex R. G. Mason [1] and Andrew J. Lowe [1]

[1] School of Biological Sciences, University of Adelaide, Adelaide 5005, Australia; nicole.foster@adelaide.edu.au (N.R.F.); greg.guerin@adelaide.edu.au (G.R.G.); a1741421@student.adelaide.edu.au (A.R.G.M.); andrew.lowe@adelaide.edu.au (A.J.L.)
[2] Department of Medical Biochemistry and Microbiology, Uppsala University, 751 23 Uppsala, Sweden
[3] Independent Researcher, Adelaide 5052, Australia; dugalugs29@gmail.com
[4] College of Science and Engineering, Flinders University, Bedford Park 5042, Australia; martin.breed@flinders.edu.au
[5] Terrestrial Ecosystem Research Network (TERN), University of Adelaide, Adelaide 5005, Australia
* Correspondence: Colette.blyth@adelaide.edu.au (C.B.); matthew.christmas@imbim.uu.se (M.J.C.)

Abstract: *Olearia pannosa* is a plant species listed as vulnerable in Australia. Two subspecies are currently recognised (*O. pannosa* subsp. *pannosa* (silver daisy) and *O. pannosa* subsp. *cardiophylla* (velvet daisy)), which have overlapping ranges but distinct leaf shape. Remnant populations face threats from habitat fragmentation and climate change. We analysed range-wide genomic data and leaf shape variation to assess population diversity and divergence and to inform conservation management strategies. We detected three distinct genetic groupings and a likely cryptic species. Samples identified as *O. pannosa* subsp. *cardiophylla* from the Flinders Ranges in South Australia were genetically distinct from all other samples and likely form a separate, range-restricted species. Remaining samples formed two genetic clusters, which aligned with leaf shape differences but not fully with current subspecies classifications. Levels of genetic diversity and inbreeding differed between the three genetic groups, suggesting each requires a separate management strategy. Additionally, we tested for associations between genetic and environmental variation and carried out habitat suitability modelling for *O. pannosa* subsp. *pannosa* populations. We found mean annual maximum temperature explained a significant proportion of genomic variance. Habitat suitability modelling identified mean summer maximum temperature, precipitation seasonality and mean annual rainfall as constraints on the distribution of *O. pannosa* subsp. *pannosa*, highlighting increasing aridity as a threat for populations located near suitability thresholds. Our results suggest maximum temperature is an important agent of selection on *O. pannosa* subsp. *pannosa* and should be considered in conservation strategies. We recommend taxonomic revision of *O. pannosa* and provide conservation management recommendations.

Keywords: *Olearia pannosa*; conservation genomics; habitat suitability; genetic diversity; gene flow; cryptic species

Citation: Blyth, C.; Christmas, M.J.; Bickerton, D.C.; Breed, M.F.; Foster, N.R.; Guerin, G.R.; Mason, A.R.G.; Lowe, A.J. Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy. *Life* **2021**, *11*, 553. https://doi.org/10.3390/life11060553

Academic Editors: Daria Sanna, Marco Casu and Fabio Scarpa

Received: 20 May 2021
Accepted: 9 June 2021
Published: 11 June 2021

1. Introduction

Populations of threatened plant species are typically small and fragmented, suffering low genetic diversity associated with increased genetic drift and elevated inbreeding rates [1–5]. Additional to these factors, a changing climate places increasing pressure on threatened populations. Plants can respond in one of three ways in the face of climate change: adapt to the new environment, migrate to locations with suitable conditions, or perish [6,7]. In this era of accelerated climate change, plants may not have the capacity to respond through in-situ adaptation, particularly those already suffering loss of genetic

diversity [8–12]. Furthermore, the sheer rate of climate change combined with limitations to migration and dispersal means that for many plant populations the likelihood of migrating at a suitable pace is low [13].

For plants, natural dispersal generally occurs via distribution of seed and pollen. For threatened plant species, however, undirected natural dispersal is likely insufficient to ensure persistence of, and gene flow among, populations [14]. In these scenarios, conservation interventions, such as facilitated gene flow, are now being considered to minimise extinction risks [15,16]. Various approaches to seed sourcing can be used, such as facilitating gene flow between fragmented populations to maximise overall genetic diversity, increasing gene flow to designated populations (e.g., to encourage adaptation to predicted climate change scenarios), and sourcing seed entirely from non-local provenances where the populations are already presumed to be well adapted to predicted future climate [1,16–23]. However even with such interventions, environmental change may take place more quickly than a species' capacity to adapt. In extreme scenarios, habitat suitability is expected to change considerably under climate change, with predications for some species showing minimal overlap between current and future suitable habitat [24,25]. In such instances, when extinction risk is high and unlikely to be alleviated by facilitated gene flow alone, establishing new populations through assisted migration (moving a species to new, more suitable habitat) and range expansion strategies can provide a last resort opportunity for conservation management [26–28].

The genus *Olearia* (Moench) is widespread across Australasia and New Guinea and there are roughly 130 species of Olearia in Australia, making it the country's largest Asteraceae genus [29]. However, many species within this genus require further investigation into their taxonomy and phylogeny. This is highlighted by the number of recent studies describing or reinstating Olearia species from both molecular and taxonomic data [29–34]. *Olearia pannosa* (Hook) is a long-lived understory daisy shrub, endemic to southeastern Australia. It generally occurs in small and fragmented stands in areas that have been converted to agriculture over the last 150 years and is classed as nationally vulnerable [34]. Significant investments have been made to conserve the species, and one of the priority conservation actions is to establish new translocated populations resilient to climate change [35].

Olearia pannosa has undergone several taxonomic revisions [34]. Currently, two subspecies are described based on leaf shape differences: *O. pannosa* subsp. *pannosa* (Cook), which is widespread across southern South Australia and has small, narrow leaves; and *O. pannosa* subsp. *cardiophylla* (F. Muell) which has larger, more heart-shaped leaves and is only found in isolated stands in the Flinders Ranges of South Australia (where it overlaps with *O. pannosa* subsp. *pannosa*) and in several isolated populations in Victoria (Figure 1). Historically, *O. pannosa* subsp. *cardiophylla* was considered a separate species and was even once within a different genus [34]. Currently there is no information on the genetic relatedness between *O. pannosa* subsp. *pannosa* and subsp. *cardiophylla*, nor on levels of genetic divergence between the Flinders Ranges *O. pannosa* subsp. *cardiophylla* populations and those in the southeast of South Australian and Victoria. Furthermore, there is a lack of understanding of the relationship between genetic divergence and leaf shape variation, which could assist practitioners with management outcomes. Therefore, to delineate a suitable conservation strategy for *O. pannosa*, we first need to understand the genetic and leaf shape differences between the various populations and subspecies.

Here, we adopt a multidisciplinary approach that combines genome-wide population genetic marker data, leaf shape and habitat suitability modelling to obtain a high-resolution picture of population structure, divergence, and diversity within this species. We aimed to test for genetic and leaf shape differences among the recognised subspecies and clarify subspecies boundaries. In particular, we aimed to reveal the relationship between isolated *O. pannosa* subsp. *cardiophylla* collection localities in the Flinders Ranges of semi-arid South Australia and those located further east. For *O. pannosa* subsp. *pannosa*. We also sought to assess whether any of the genetic and phenotypic diversity among the collection localities

associates with environment and used habitat suitability modelling to reveal environmental limitations to the species' distribution. We discuss the outcomes from these analyses in terms of their implications for the conservation management.

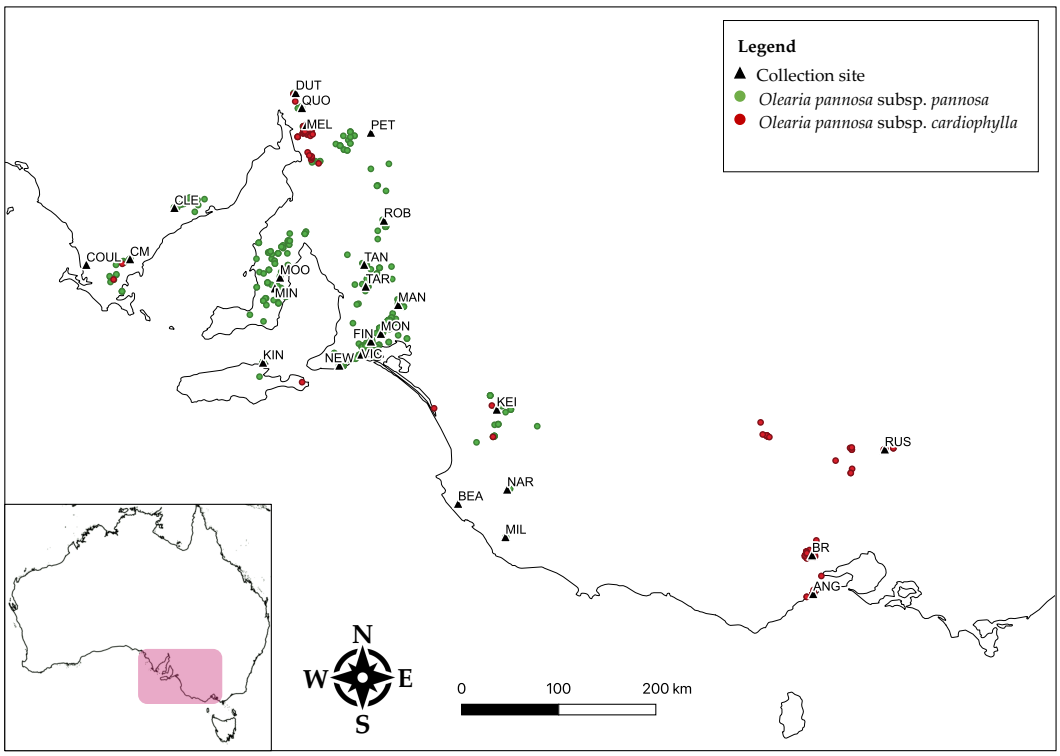


Figure 1. Study collection sites (black triangles) and occurrence records of *O. pannosa* subsp. *pannosa* (green markers) and *O. pannosa* subsp. *cardiophylla* (red makers) (Atlas of Living Australia occurrence downloads at https://biocache.ala.org.au/occurrences/search?q=qid:1609477224810 and https://biocache.ala.org.au/occurrences/search?q=qid:1609477191724, both accessed on 1 January 2021).

2. Materials and Methods

2.1. Study Species and Sampling

Olearia pannosa is a low shrub (~1.5 m) which spreads by producing root suckers [34]. The species can be found in a range of floristic communities and climatic conditions within southeastern Australia. There are currently two subspecies with similar ecology but can be differentiated by leaf shape: *O. pannosa* subsp. *pannosa* (silver daisy bush), and *O. pannosa* subsp. *cardiophylla* (velvet daisy bush). *O. pannosa* subsp. *pannosa* is listed as vulnerable under the Australian Government's Environment Protection and Biodiversity Conservation Act 1999 [36]; and vulnerable under the South Australian National Parks and Wildlife Act 1972 [37]. *O. pannosa* subsp. cardiophylla is also of conservation concern and has been listed as vulnerable under the Victorian Advisory List of Rare or Threatened Plants in Victoria—2014 [38].

2.2. Study Species and Sampling

Leaf samples were collected from a total of 191 individuals from 25 localities spanning the ranges of both subspecies (Figure 1, Table A1). Each sample (individual plant) consisted of five leaves: one leaf collected from each aspect/side of the plant (North, South, East, West), and one from the central apex point of the plant—generally the highest/near highest

growing point of the plant. Nearest neighbour individuals were avoided to minimise relatedness between genotyped individuals and obvious clones and root suckers were avoided. Samples were dried in a sealed container with silica beads. Collections were made during winter and spring of 2018. Voucher specimens were also collected at each locality (except the Victorian stands), and where there was evidence of morphological variation between individuals at a site, multiple specimens were collected. Voucher specimens consisted of a 20–30 cm small branch taken from a growing point of the plant, which was labelled and pressed at the earliest convenience (Figure A1). Vouchers were submitted to the State Herbarium of South Australia, and an identification was provided for each collection locality (Table A1).

2.3. DNA Extraction, Sequencing, and Filtering

All DNA extraction, library preparation and sequencing were conducted at the Australian Genome Research Facility (Adelaide, Australia). The Machery-Nagel Nucleospin Plant II Kit was used to extract DNA prior to undergoing double digest restriction-associated DNA sequencing (ddRADseq) [39]. The restriction enzymes Pstl and MseI were used for the digestion, then barcoded adaptors were ligated to the restriction site overhangs for each sample. Samples were then pooled, and the DNA size selected using Blue Pippin (Sage Science, Beverly, MA, USA) (Wide) to retain fragments 280–342 bp. The resulting library was amplified via Polymerase Chain Reaction (PCR) for 11 cycles using indexed primers. Libraries were assessed by gel electrophoresis (D1000 ScreenTape Assay, Agilent, CA, USA), quantified by qPCR (KAPA Library Quantification Kits for Illumina, Roche, Basel, Switzerland), and then single read sequenced (150 bp) on the NextSeq 500 system using NextSeq 500 high Output Kit v2 reagents (150 cycles).

A total of 637,395,304 sequences were generated across 192 *O. pannosa* samples and we processed these using the STACKS pipeline [40,41] at AGRF (Melbourne, Australia). Reads were de-multiplexed using the unique sample adapter barcodes and filtered for both read quality and the presence of a restriction site. The resulting FASTQ files were then trimmed to the size of the shortest read, minus 2 bp to account for differences in read length due to any variation in barcode length. Reads were then stacked according to similarity, forming tags of reads. Tags which appear across all samples were collated (catalogue tags), and genotypes were then allocated to the common polymorphic sites. The collated SNPs across all individuals were then filtered using the following settings: minimum number of reads required at a stack to call a homozygous genotype = 5; minor allele frequency, below which a stack is called a homozygote = 0.05; minimum minor allele frequency to call a heterozygote = 0.1 (above 0.05 but below 0.1, a stack is called 'unknown'). Computed genotypes were then output as a single VCF file (Supplementary File S1) for further filtering and analysis.

We used VCFtools [42] to filter SNPs with a minor allele frequency of less than 0.1 and with a depth of coverage of less than10 reads per individual. We removed SNPs with >25% missing data across all individuals and individuals with >25% missing SNP calls. We then used PLINK [43] to identify SNPs in linkage disequilibrium with the 'indep-pairwise' command, with one SNP in a pair of linked SNPs ($r^2 > 0.8$) being removed from the dataset.

Levels of ploidy were assessed by examining the distribution of allele dosage at heterozygous sites. For this, the counts of reads supporting each allele at a heterozygous site were extracted from the VCF file using a custom Perl script, and each individual's distribution was plotted. True allelic proportions for diploids, triploids and tetraploids at heterozygous sites are 1:1 (0.5), 1:2 or 2:1 (0.33 or 0.66), and 1:3, 2:2, or 3:1 (0.25, 0.5, or 0.75), respectively. No sample deviated from a normal distribution centred on a frequency of 0.5, suggesting that all samples are diploid.

For each SNP across all samples, inbreeding coefficient (F) values were estimated using the G-stats calculator in GENODIVE v. 3.2 [44,45]. As negative F indicates greater than expected heterozygosity under Hardy-Weinberg equilibrium, and could be indicative of paralogous reads being stacked together in the STACKS pipeline [46], we used VCFtools

to remove SNPs with significantly negative F (significance assessed using permutation tests with 10,000 permutations and a significance threshold of $p < 0.05$).

After the filtering steps, 175 individuals and 9997 SNPs remained. We used this dataset to perform the PCA, DAPC, and ADMIXTURE. After filtering for highly related individuals, 169 individuals and 9997 SNPs remained in the dataset. We used this dataset to perform F_{ST}, genetic diversity and inbreeding statistics, and AMOVA analyses.

2.4. Genetic Structure Analysis

We then analysed the population genetic structure across all samples in the filtered SNP. To identify the most likely number of genetic clusters (K), we used two different genetic clustering analyses. First, we used the non-model-based Discriminant Analysis of Principal Components (DAPC) of the R package adegenet [47]: genetic data were transformed into uncorrelated components using principal component analysis (PCA). The number of genetic clusters was then defined using k-means, a clustering algorithm that looks for the value of K that maximises the variation between groups. The Bayesian Information Criterion (BIC) was calculated for K = 1–20, and the K value with the lowest BIC suggesting the optimal number of clusters. A discriminant analysis was then performed on the first 80 principal components with the function 'dapc', implemented in R, in order to efficiently describe the genetic clusters and assign samples to each cluster. We also used ADMIXTURE [48], a model based genetic clustering algorithm, to estimate the most likely number of clusters (K) in our dataset. ADMIXTURE was also run for K values 1–10 and the most likely value of K was assessed by comparing the cross-validation errors (cv errors) between runs, with the lowest cv error indicating greatest support. We repeated each ADMIXTURE analysis ten times and used CLUMPAK to combine the results and construct bar plots of individual assignment to clusters for the most supported values of K using DISTRUCT [49,50]. We then calculated pairwise F_{ST} between all collection localities using the R package StamPP [51]. This was then exported to Splitstree to create a neighbour-joining (NJ) tree [52].

2.5. Kinship, Genetic Diversity, and Inbreeding Analysis

The dataset was filtered for all first-degree relations (parent-siblings and sibling-siblings) following methods outlined in Dutheil [53] to avoid these influencing our statistical results. The king-cutoff option in PLINK 2 [54] was used to remove one in a pair of samples with a kinship coefficient greater than or equal to 0.177 (6 individuals were removed, Tables A2 and A4). A kinship coefficient of 0.25 corresponds to first-degree relations, so the 0.177 threshold should capture all of these. This step should also have removed any likely clones from the dataset, which was an important consideration given *O. pannosa* is known to reproduce clonally. We then used PLINK 2 [54] to calculate pairwise kinship between individuals. We then calculated average kinship within each locality in R v.4.0.4 As the genetic structure analysis identified such strong genetic differentiation between three of the clusters, we separated the dataset according to these groupings. Each dataset was refiltered for minor allele counts of 1 with VCFtools [39] to remove non-variants for subsequent analyses (See Table A2 for number of variants remaining at each collection locality).

For the collection localities in each of the three genetic groups, we separately estimated the inbreeding coefficient (F), and observed (H_O) and expected (H_E) heterozygosity in GENODIVE v3.2 [44,45]. For the collection locality COUL (COULTA), there were insufficient samples to calculate these statistics and so this collection locality was not included in this analysis. A one-way ANOVA and Tukey pairwise comparisons were used to assess group differences in kinship, F, H_O and H_E using R v.4.0.4 [55]. To assess the distribution of genetic diversity amongst and within the genetic groups, we conducted a nested Analysis of Molecular Variance (AMOVA) using GENODIVE v3.2 to explore genetic variation within individuals, among collection localities nested within groups, and among groups. Further AMOVA analyses were then run separately for *O. pannosa* subsp. *pannosa* and *O. pannosa*

subsp. *cardiophylla* (there was not enough data to complete the analysis for *Olearia* (FR)) to compare how genetic variation was distributed within individuals and among collection localities between these two genetic groups.

2.6. Leaf Trait Analysis

In order to investigate the environmental influence on leaf shape and the variation that exists between collection localities, leaf traits were identified and analysed. In total 110 individuals were measured from across 22 collection localities (note that Victorian collection localities were excluded from this component of the study as we did not have suitable material for the measurements). Leaf traits were based on the methods of Lowe and Abbott [56] and measured using ImageJ software v 1.53e (https://imagej.nih.gov/ij/download.html, accessed on 10 November 2020) [57] from photo images of voucher samples (See Figure A1 for examples). Three measurements were taken; leaf length (L) (measured from leaf base to apex), leaf width (W) (measured at the widest point of leaf), and the distance along the midrib to the widest point of the leaf from the base of the leaf, giving width location (WL). Leaf area measurements were square root transformed to satisfy the requirement of normality of residuals. Leaf length (L), width (W) and the distance at which the width measurement was taken from the base (WL) were used to estimate leaf area (LA), employing the formula

$$LA = (L \times W / 2) + ((L - WL) \times 2 / 2) \quad (1)$$

We estimated leaf ovality or roundness by dividing the length along leaf mid rib to widest point of the leaf by total leaf length. A value around 0.5 indicates an oval leaf, <0.5 indicates a more heart-shaped leaf, and >0.5 a 'gingko' shaped leaf. The ratio of leaf length to leaf width (calculated as leaf length divided by width), was also determined. Where possible, five leaves were measured from each voucher specimen, with five individuals sampled from collection locality. A standardised leaf sampling protocol was conducted by sampling the first mature leaf from the base of the specimen, with subsequent leaves sampled working towards the apex. In some circumstances direct leaf measurement was not possible due to folding or damage, in which case width was calculated by measuring distance from leaf edge (at widest point) to leaf midrib and multiplying by two.

A principal component analysis (PCA) was performed using the 'prcomp' function in R v.4.0.4 [55] on the measurements for all measured leaf samples to assess differences in leaf shape among the three genetic clusters. Trait values were centred and scaled. One-way ANOVA and Tukey pairwise comparisons were used to assess group differences in principal component axes 1 (PC1) and 2 (PC2) using R v.4.0.4 [55].

2.7. Redundancy Analyses—Both Genetic Data and Leaf Variation

To assess the degree to which the environment explains observed genetic and phenotypic variation, partial redundancy analyses (RDA) were performed for *O. pannosa* subsp. *pannosa* samples. We focused on this genetic group as they were distributed over a wide environmental gradient and by omitting *O. pannosa* subsp *cardiophylla* samples confounding genetic structure can be removed from the analysis. A set of 44 environmental (e.g., climate, soil, and landscape) variables were sourced from stacks of 9-s resolution (i.e., approximately 250 m pixels) rasters [58–60]. Environmental values were extracted for each raster at the coordinates of collection localities of *O. pannosa* subsp. *pannosa* using the RASTER analytical package in R [61]. A subset of these extracted environmental variables with low collinearity were assessed using the 'cor' function in R v.4.0.4 [55]. They were: mean annual maximum temperature (°C), mean annual minimum temperature (°C), annual precipitation (mm), total nitrogen (%), total phosphorus (%), organic carbon (%), and elevation (m).

Redundancy analyses (RDA) were performed on (a) the allele frequencies of all SNPs in collection localities of *O. pannosa* subsp. *pannosa* and (b) leaf trait measurements for all samples of *O. pannosa* subsp. *pannosa* in order to assess how much of the genetic

and phenotypic variation among our collection localities can be explained by spatial (i.e., resulting from isolation by distance) and environmental (i.e., indicating selection) factors (Supplementary File S2) using the VEGAN analytical package v2.5-6 in R [62]. First, spatial coordinates for (a) each sampling stand or (b) each sample were centred and converted into third degree polynomials. An RDA of allele frequencies-all spatial polynomials was then performed using the 'rda' function. The 'ordistep' function was then used to carry out forward stepwise model building for the RDA, adding in one polynomial to the model at a time and assessing significance using permutation tests. This revealed x, y, and x^2 to be significant contributors to the model and as such only these variables were retained for the RDA. The variance between the selected spatial and all environmental variables was assessed using the 'varpart' function and significance of the partitioning was assessed using an analysis of variance (ANOVA)-like permutation test for RDA with the 'anova.cca' function. In each case, we tested whether any of the genetic or phenotypic variance was significantly explained by environmental variables by running an RDA on the models (a) 'allele frequencies-environmental variables conditioned on spatial variables' and (b) 'leaf traits-environmental variables conditioned on spatial variables'. To assess which of the environmental variables provided the most explanatory power in the models, we performed the same stepwise model building described above for the spatial variables on the environmental variables.

2.8. Habitat Suitability Modelling

For the largest genetic group, *O. pannosa* subsp. *pannosa*, MaxEnt and Random Forest were used to assess which environmental variables best explain habitat suitability given the species' current distribution. Occurrence data were sourced from collection locality sampling records, consisting of 105 locations. We modelled current distribution using the BCCVL online tool (http://www.bccvl.org.au/, accessed on 25 February 2021); Ref. [63] by statistically comparing confirmed presence sites against predictor spatial layers. Predictors included national soil data provided by the Australian Collaborative Land Evaluation Program ACLEP (www.clw.csiro.au/aclep, accessed on 25 February 2021), and a current Australian climate baseline of 1976 to 2005 provided as bioclim layers by BCCVL [64].

Distributions were predicted based on current climate and soil using MaxEnt and Random Forest, which are widely used machine-learning algorithms for mapping habitat suitability [65,66]. Models were run using 10,000 pseudo-absences confined to the land area within 250 km of a convex polygon surrounding presence locations. Starting with the full set of potential variables included in the edaphic and climatic layer sets, we sequentially removed both highly correlated (i.e., $r > 0.8$) and poorly performing predictors contributing to 1% or less of the model.

3. Results

3.1. Genetic Structure Analysis

Strong genetic structure was identified across the range of *O. pannosa* and was supported by all genetic analyses (Figures 2, 3, A2 and A3, Table A3). While the BIC criterion from the DAPC analysis (Figure A3) and CV error from the ADMIXTURE (Figure A2) both indicate that K = 5–6 is the most likely value of K, there is strong evidence of three key genetic groups within *O. pannosa* which have genetic barriers that cannot be explained by geographic isolation (Figure 2). This was shown in samples from the Flinders Ranges in semi-arid South Australia identified as *O. pannosa* subsp. *cardiophylla*. These were genetically distinct from all other samples, even when collected from the same geographic location (see collection site Melrose (MEL) in Figures 2, 3, A1 and A2, Table A3). The second such genetic break is found on the Fleurieu Peninsula of South Australia. Despite evidence of admixture in these collection localities (Figure 2), the geographically proximate Newland Head (NEW) and Victor Harbor (VIC) (~12 km apart) fall into different clusters in the PCA and different genetic groups in the DAPC analysis (Figures 2, 3 and A3, Tables A1 and A3).

We therefore present three key genetic groups, acknowledging that there is further substructure within them. These groups are:

1. The southern Flinders Ranges samples identified by the State Herbarium of South Australia as *O. pannosa* subsp. *cardiophylla* (Table A1, Peter Lang, pers. comm.). From here on, we will refer to this group as *Olearia* (FR).
2. Samples from Coulta (COUL) on the Eyre Peninsula to Victor Harbor (VIC) on the Fleurieu Peninsula (excluding the *Olearia* (*FR*) specimens collected at Dutchman's Stern (DUT) and Melrose (MEL)) were mainly identified as *O. pannosa* subsp. *pannosa* (State Herbarium of South Australia, P. Lang 2021, personal communication; Table S1). There is evidence of some substructure within this group, with geographically isolated sites Coulta (COUL) and Cummins (CM) on the Eyre Peninsula forming a distinct cluster. Notably, there was some admixture between genetic groups, particularly in the samples collected at Victor Harbor (VIC), which were identified as *O. pannosa* subsp. *cardiophylla* (Table A1). From here on this genetic group will be referred to as *O. pannosa* subsp. *pannosa.*

All other samples ranging from the NEW (Newland Head) collection locality on the Fleurieu Peninsula across to the Victorian collection localities were identified as *O. pannosa* subsp. *cardiophylla* (State Herbarium of South Australia, P. Lang 2021, personal communication; Table A1). However, admixture was also detected in several collection localities geographically close to stands of *O. pannosa* subsp. *pannosa* (e.g., Newland Head (NEW)). From hereon we will call this genetic group *O. pannosa* subsp. *cardiophylla.*

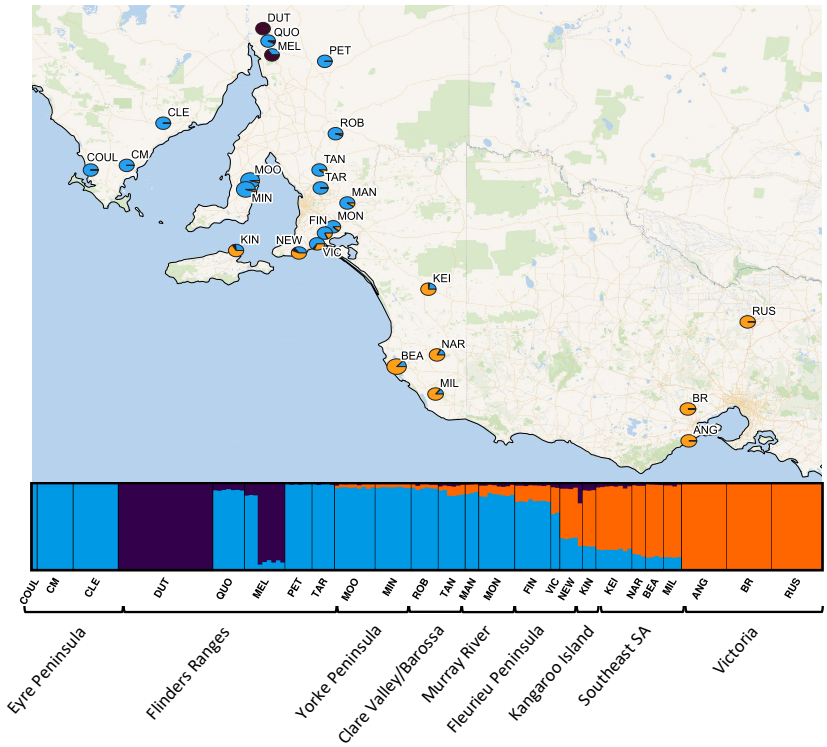


Figure 2. Individual genetic cluster assignment based on ADMIXTURE results. Pie charts at each sampling site show the overall proportion of individuals assigned to the two genetic clusters (blue generally aligns with collection localities identified as *O. pannosa* subsp. *pannosa* group, purple with *Olearia* (FR), and orange with *O. pannosa* subsp. *cardiophylla.* Individual assignments are represented in the bar plots at the bottom of the map.

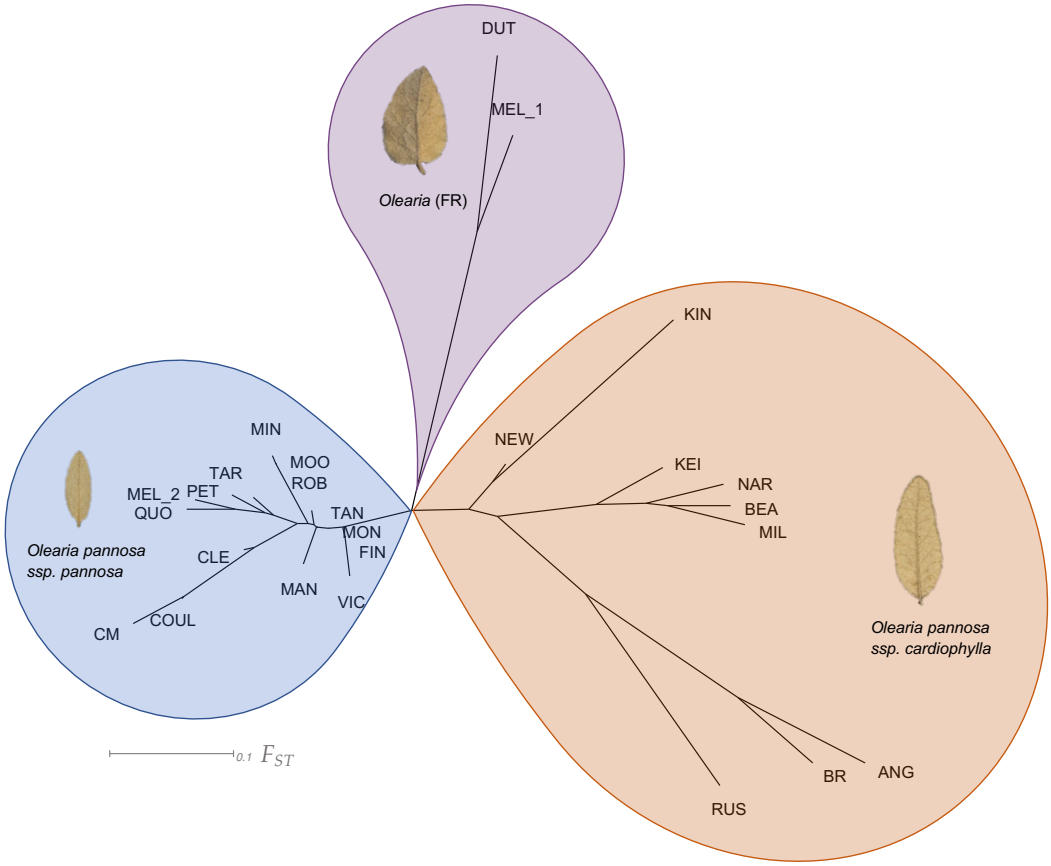


Figure 3. Neighbour joining tree of pairwise genetic differentiation (F_{ST}) between collection localities. Colours represent genetic cluster assignment identified in the ADMIXTURE and DAPC analyses (blue: collection localities identified as *O. pannosa* subsp. *pannosa* group, purple: *Olearia* (FR), orange: *O. pannosa* subsp. *cardiophylla*). Images of leaves represent typical shape found in each genetic cluster. The *O. pannosa* subsp. *cardiophylla* cluster leaf is from Naracoorte (NAR), the *Olearia* (FR) cluster from Melrose (MEL), and the *O. pannosa* subsp. *pannosa* cluster leaf is also from Melrose (MEL). Note that as the site Melrose (MEL) contains individuals belonging to different genetic clusters, those belonging to *Olearia* (FR) have been labelled as MEL_1 and *O. pannosa.* subsp. *pannosa* have been labelled as MEL_2 (see Table A3 for full output and Figure A1 for images of full voucher specimens).

Across the whole dataset, our analysis of pairwise genetic differentiation (F_{ST}) showed a considerable range of differentiation between collection localities (F_{ST} = 0.00–0.75; Figure 3, Table A3). Within the *Olearia* (FR) group (DUT & MEL_1), there was a moderately high level of genetic differentiation (F_{ST} = 0.23). Within *O. pannosa* subsp. *pannosa*, apart from collection localities on the Eyre Peninsula, genetic differentiation between collection localities was generally low (F_{ST} = 0.00–0.13). On the Eyre Peninsula, however, collection localities displayed a greater range of differentiation (F_{ST} = 0.05–0.30) reflecting their more isolated locations. There was considerable variation in the levels of genetic differentiation between collection localities in *O. pannosa* subsp. *cardiophylla* (F_{ST} = 0.12–0.68). Notably, the small and isolated Kangaroo Island collection locality at Kingscote (KIN), had disproportionately higher levels of F_{ST} compared to all other collection localities (F_{ST} = 0.36–0.68). Collection localities within the Fleurieu Peninsula and southeast SA, show moderate differentiation (F_{ST} = 0.17–0.22) but are highly differentiated from the geographically distant and isolated Victorian collection localities (F_{ST} = 0.36–0.59). The collection localities within the Victorian

region showed moderate genetic differentiation among Anglesea (ANG) and Brisbane Ranges (BR) (F_{ST} = 0.20) but high when compared to the most isolated and northern Victorian locality Rushworth (RUS) (F_{ST} = 0.46 and 0.42 respectively).

3.2. Genetic Diversity, Kinship, and Inbreeding Analysis

Across the dataset, the majority of genetic variation was found within individuals, but that a significant proportion of variation was also distributed between the three genetic groups (AMOVA; 42.7% and 29.9% respectively; Table 1), however there were differences between subspecies. Within *O. pannosa* subsp. *pannosa*, most genetic variation was found within individuals (67.8%) and only 11.2% between collection localities (Table 1). However, within *O. pannosa* subsp. *cardiophylla* a high proportion of the genetic variation was distributed between collection localities (43.2%; Table 1).

Across all samples and groupings, observed heterozygosity (H_O = 0.096–0.245) was consistently lower than expected heterozygosity (H_E = 0.109–0.330; Figure 4, Table S2), and positive inbreeding coefficients (*F*) were found in all localities (*F* = 0.055–0.325; Figure 4, Table A2). Kinship was negative in each locality (kinship = −0.062 – −0.499; Figure 4, Table A2).

Table 1. Nested analysis of molecular variance (AMOVA) analysing the distribution of genetic diversity (**a**) across all collection localities, (**b**) within *O. pannosa* subsp. *pannosa*, and (**c**) within *O. pannosa* subsp. *cardiophylla*. Including standard deviations (SD; obtained through jackknifing over loci) and 95% confidence intervals (c.i. 95%; obtained through bootstrapping over loci) of F statistics.

(a) All collection localities							
Source of variation	Nested in	%var	F	Std.Dev.	c.i.2.5%	c.i.97.5%	*p*
Within individual	–	42.7	0.573	0.002	0.569	0.578	–
Among individuals	Collection locality	11.2	0.208	0.003	0.203	0.213	<0.001
Among collection localities	group	16.2	0.231	0.001	0.228	0.234	<0.001
Among group	–	29.9	0.299	0.002	0.294	0.304	<0.001
(b) *Olearia pannosa* subsp. *pannosa* only							
Source of variation	Nested in	%var	F	Std.Dev.	c.i.2.5%	c.i.97.5%	*p*
Within individual	–	67.8	0.322	0.002	0.317	0.327	–
Among individuals	Collection locality	21.3	0.239	0.003	0.234	0.244	<0.001
Among collection localities	–	10.9	0.109	0.001	0.107	0.111	<0.001
(c) *Olearia pannosa* subsp. *cardiophylla* only							
Source of variation	Nested in	%var	F	Std.Dev.	c.i.2.5%	c.i.97.5%	*p*
Within individual	–	48.2	0.518	0.003	0.512	0.525	–
Among individuals	Collection locality	8.6	0.152	0.004	0.144	0.16	<0.001
Among collection localities	–	43.2	0.432	0.003	0.427	0.438	<0.001

Within *Olearia* (FR), genetic diversity was low (H_O = 0.164–0.170; Figure 4, Table A2), historical inbreeding was high (*F* = 0.241–0.332; Figure 4, Table A2), but levels of kinship were low to moderate (kinship = −0.388 – −0.216; Figure 4, Table A2). *Olearia* (FR) samples collected at Melrose (MEL_1) had highest inbreeding (*F* = 0.332) but lower levels of kinship (kinship = −0.388). Within *O. pannosa* subsp. *pannosa*, levels of genetic diversity (H_O = 0.174–0.245; Figure 4, Table A2) and inbreeding (*F* = 0.211–0.330; Figure 4, Table A2) did not vary much between localities. There was greater variation in kinship (kinship = −0.476–−0.242; Figure 4, Table A2), with the admixed locality Victor Harbor (Vic) having the lowest levels of kinship (kinship = −0.476; Figure 4, Table A2). There is a pattern of lower kinship in the localities with higher inbreeding. There is also a general trend of expected heterozygosity (H_E) and observed heterozygosity (H_O) increasing the closer collection localities are to the Fleurieu Peninsula in southeast South Australia, align with increased levels of admixture detected in this region (Figure 4, Table A2). Within

O. pannosa subsp. *cardiophylla*, Newland Head (NEW) and Keith (KEI) had the highest levels of genetic diversity and inbreeding (H_E = 0.315 and 0.239, H_O = 0.224 and 0.172, F = 0.288 and 0.279, respectively; Figure 4, Table A2), and lowest levels of kinship (kinship = −0499 and −0.350, respectively; Figure 4, Table A2). All other localities had moderate expected heterozygosity (H_E = 0.09–0.190; Figure 4, Table A2). lower inbreeding coefficients than *O. pannosa* subsp. *pannosa* (F = 0.08–0.17; Figure 4, Table A2), but high levels of kinship (kinship = −0.213 – −0.062; Figure 4, Table A2). Furthermore, *O. pannosa* subsp. *cardiophylla* was the only group to have highly related individuals removed; Brisbane Ranges (BR), Kangaroo Island (KIN), and Anglesea (ANG).

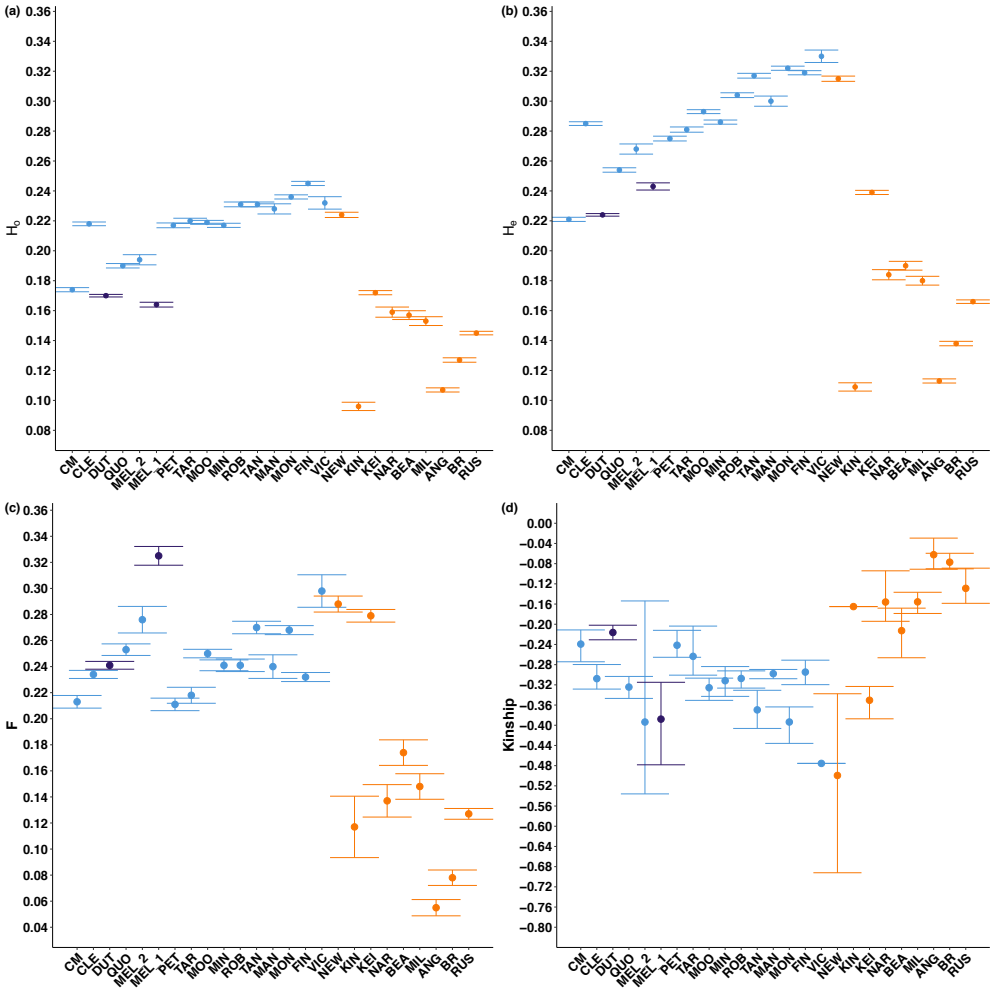


Figure 4. Scatterplots of (**a**) observed heterozygosity (H_O), (**b**) expected heterozygosity (H_E) (both vary from 0 to 1, with 0 indicating no heterozygous sites), (**c**) inbreeding coefficient (F) (varies from 0 to 1 and values of 0 indicate no evidence of inbreeding), and (**d**) kinship (0.5 indicates individuals are genetically identical). Bars represent 95% confidence intervals. Colours represent genetic group (blue with collection localities predominantly identified as *O. pannosa* subsp. *pannosa*, purple for *Olearia* (FR), and orange for *O. pannosa* subsp. *cardiophylla*). See Table S2 for full output. For collection locality full names refer to Table A1. For results of the Tukey pairwise comparison refer to Table S5.

The ANOVA analysis and Tukey pairwise comparisons of observed heterozygosity (H_O), expected heterozygosity (H_E), inbreeding (F), and kinship, revealed significant differences for *O. pannosa* subsp. *pannosa* and *O. pannosa* subsp. *cardiophylla* ($p < 0.001$; Table A5). There was no significant difference in Ho, He, F, and kinship for *Olearia* (FR) compared to *O. pannosa* subsp. *pannosa* ($p > 0.05$; Table A5). The only significant difference between *Olearia* (FR) and *O. pannosa* subsp. *cardiophylla* was inbreeding coefficient (F) in *Olearia* (FR) ($p = 0.02$; Table A5).

3.3. Leaf Trait Analysis

Principal component analysis of all leaf measurements revealed clear clustering of samples, which broadly aligned with the three genetic groups, but with some overlap, indicating morphological intergrades (Figure 5). Principal component 1 (PC1) explained 68.5% of the variance and mainly explained by variation in length, width, and surface area of leaves (Figure A5). *Olearia pannosa* subsp. *pannosa* separated from all other samples on axis 1 due to their shorter, smaller leaves. Principal component 2 (PC2) explained 25.6% of the variance and mainly explained variation in leaf ovality and location of the widest point of the leaf. On this axis, the majority of *Olearia* (FR) samples were separated from *O. pannosa* subsp. *cardiophylla* and *O. pannosa* subsp. *pannosa* samples due to their wider, more heart-shaped leaves. For PC1 (width, length, and leaf surface area), our ANOVA analysis found significant differences between *O. pannosa* subsp. *pannosa* and *Olearia* (FR)/*O. pannosa* subsp. *cardiophylla* ($p < 0.001$; Table A6). For PC2 (ovality and location of the widest point of the leaf), differences between *O. pannosa* subsp. *pannosa* and *O. pannosa* subsp. *cardiophylla* were not significant, but both groups differed significantly to *Olearia* (FR) ($p < 0.001$; Table A6).

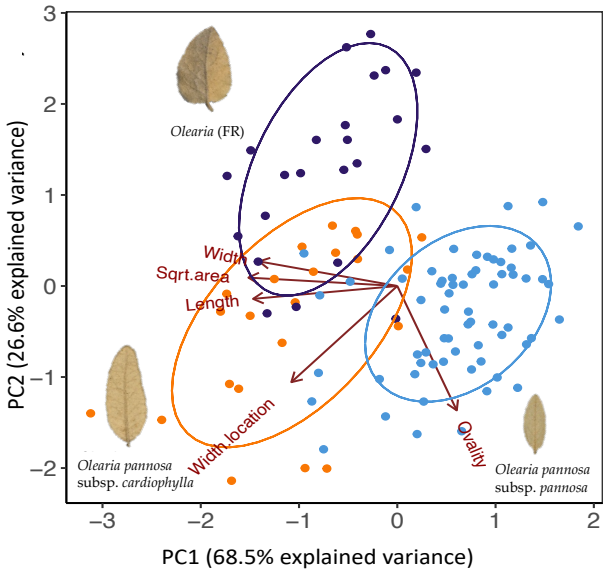


Figure 5. PCA of leaf shape variation. Colours represent genetic cluster assignment identified in the ADMIXTURE and DAPC analyses (blue: collection localities identified as *O. pannosa* subsp. *pannosa* group, purple: *Olearia* (FR), orange: *O. pannosa* subsp. *cardiophylla*). Images of leaves represent the typical shape found in non-admixed populations of each genetic cluster. The *O. pannosa* subsp. *cardiophylla* cluster leaf is from Naracoorte (NAR), the *Olearia* (FR) cluster from (MEL), and the *O. pannosa* subsp. *pannosa* cluster leaf is also from Melrose (MEL). Note that as the site Melrose (MEL) contains individuals belonging different genetic clusters. Those belonging to *Olearia* (FR) have been labelled as MEL_1 and *O. pannosa*. subsp. *pannosa* have been labelled as MEL_2.

3.4. Environmental Associations

The redundancy analysis of *O. pannosa* subsp. *pannosa* collection localities revealed that spatial distribution explained a significant proportion of the variance in allele frequencies (Table 2, Figure A4). The partial redundancy analysis, which looked at the proportion of variance explained by the seven environmental variables (mean annual maximum temperature (°C), mean annual minimum temperature (°C), annual precipitation (mm), total nitrogen (%), total phosphorus (%), organic carbon (%), and elevation (m)) after conditioning on spatial distribution, revealed that 11% of the variance in allele frequencies is significantly explained by the included environmental factors (Table 2, Figure A4). Forward selection of the environmental variables retained only mean maximum temperature as a significant contributor to the model, suggesting that this environmental variable associates with genetic variation.

The redundancy analysis of leaf shape within *O. pannosa* subsp. *pannosa* samples included the same seven environmental variables as for the allele frequency RDA. When running the RDA with only spatial distribution as the explanatory variable, 13% of the variance in leaf shape was significantly explained Table 2). With environment as the explanatory variable conditioned on spatial distribution only 3% of the variance is explained and was non-significant (Table 2). When partitioning the variance in leaf measures among the spatial and environmental variables, 17% of the variation is explained by the combination of both environment and spatial distribution. The environmental variables (particularly temperature and precipitation) do co-vary with spatial distribution, so it is difficult to tease apart the effects of the environment on leaf shape (i.e., environmentally determined) from differences in leaf shape that may be due to demography. However, as for the genotype RDA, the forward selection process revealed maximum annual temperature, as well as total nitrogen, as significant contributors to the model, once again suggesting that temperature associates with variation among these populations.

Table 2. Results from partial redundancy analyses. Significance was assessed using an ANOVA-like permutation test with 1000 permutations.

Response Variable	Explanatory Variables	Condition Variables	r^2	F	*p*
Allele frequencies	Space	-	0.36	2.59	<0.001
	Environment	Space	0.11	1.64	0.038
Leaf shape	Space	-	0.13	5.12	0.002
	Environment	Space	0.03	2.13	0.086

3.5. Habitat Suitability Modelling

Final models were fitted with two edaphic (% soil clay—clay30; % available water capacity—pawc1m), two temperature predictors (mean minimum temperature of the coldest month—bio6; mean maximum temperature of the warmest month—bio5), and three rainfall (mean annual rainfall—bio12; mean rainfall of the driest month—bio14; precipitation seasonality/coefficient of variation—bio15) predictors (Figures A5 and A6). Distribution models showed *O. pannosa* subsp. *pannosa* occurs within a consistent climatic band intermediate between the mesic areas and the more arid inland. Peak suitability was associated with mean annual rainfall of 400–500 mm (occupied niche 330–670 mm) and mean summer temperature maxima <29 °C (MaxEnt; <30 °C for Random Forest; occupied niche 25–31 °C). For MaxEnt, precipitation seasonality had the highest variable relative contribution followed by mean maximum temperature of the warmest month. For random forest, mean maximum temperature of the warmest month had the highest variable importance followed by mean annual rainfall.

4. Discussion

Our analysis of genomic and leaf shape variation in the threatened daisy bush, *Olearia pannosa*, unearthed a complex picture that is not in agreement with the current subspecific

classifications for the species. A key finding of our study was the identification of a likely cryptic species *Olearia* (FR) in the Flinders Ranges which exists only in small, fragmented collection localities. For the two subspecies, *O. pannosa* subsp. *pannosa* and *O. pannosa* subsp. *cardiophylla*, we identified a strong east/west genetic divide with signatures of admixture between the two subspecies and substructure within them. Genetic diversity was reduced in *Olearia* (FR) compared to *O. pannosa* subsp. *pannosa*, and was substantially lower in the small, isolated localities of *O. pannosa* subsp. *cardiophylla*. Our analyses of associations between environmental and genetic and phenotypic variation identified a strong association with maximum temperature. This finding was further supported by habitat suitability modelling, which identified maximum temperature and precipitation as key variables in determining species' distribution. Our results have implications for the conservation management for *O. pannosa*, particularly for the newly identified genetic group *Olearia* (FR) in the Flinders Ranges.

4.1. Genetic Structure, Diversity, Inbreeding, and Kinship

We found evidence of strong differentiation among collection localities, with clear support for three major genetic groups across all analyses and substructure with them. Previously, *O. pannosa* plants found in the southern Flinders Ranges were thought to be the same subspecies as *O. pannosa* subsp. *cardiophylla* found in southeastern South Australia and Victoria. However, our analysis reveals that these southern Flinders Ranges plants (*Olearia* (FR)), sampled at Dutchman's Stern (DUT) and Melrose (MEL), are not only strikingly distinct genetically from their neighbouring collection localities of *O. pannosa* subsp. *pannosa* but also from all other collection localities of *O. pannosa* subsp. *cardiophylla* in the southeastern South Australia and Victoria. This is in clear contrast to the current classification based on morphology, where *O. pannosa* in the Flinders Ranges is considered to be part of the same subspecies as *O. pannosa* subsp. *cardiophylla* found in Victoria. The apparent lack of gene flow among *Olearia* (FR) and neighbouring *O. pannosa* subsp. *pannosa* collection localities strongly suggests that there are reproductive barriers present between these two clusters and these collection localities likely represent two distinct species. Notably, both *Olearia* (FR) collection localities have high genetic differentiation between them, low genetic diversity, and high historical inbreeding. These findings are consistent with results from other recent genomic studies of threatened and endangered species [67–69]. As far as we know, this is a new finding, signalling the need for taxonomic description of this newly discovered genetic group and a conservation management strategy to be developed to ensure its survival.

The remaining samples could be assigned to either the *O. pannosa* subsp. *pannosa* or *O. pannosa* subsp. *cardiophylla* genetic groups, which are split by an east to west genetic divide. These two genetic groups differed significantly in levels of genetic diversity, inbreeding and kinship, with *O. pannosa* subsp. *cardiophylla* having lower levels of genetic diversity and inbreeding but higher levels of kinship. Interestingly, a self-compatibility test was undertaken at the Brisbane Ranges locality which found that individuals could self-pollinate and set seed, and inbreeding was flagged as a concern in *O. pannosa* subsp. *cardiophylla* [34]. Certainly *O. pannosa* subsp. *cardiophylla* shows signs of population genetic decline indicated by the low genetic diversity and high kinship, and our findings consistent with the fact that the Victorian populations are now small and demographically isolated. There is also potential for elevated levels of clonality in these collection localities [70,71], but as our sampling protocol was designed to minimise the likely number of clones in our dataset, it is not possible for us to calculate the proportion of clones within collection localities. We did, however, detect one likely clone in the Kangaroo Island (KIN) collection locality, which also had the lowest levels of genetic diversity. Nevertheless, due to the limitations in our study, we advise further research to test mating system dynamics.

We observed genetic admixture on the southern Fleurieu Peninsula and southeastern SA, with a shift from a collection locality with high assignment to the *O. pannosa* subsp. *pannosa* group at Victor Harbor (VIC) to a collection locality with high assignment to the

O. pannosa subsp. *cardiophylla* group at the closely located (~12 km) collection locality at Newland Head (NEW). Interestingly, Newland Head (NEW) had the variation in pairwise kinship between individuals within *O. pannosa* subsp. *cardiophylla.* Genetic admixture patterns were reflected in the leaf shape analysis, with intermediate morphologies described at these sites. Distribution records for both subspecies do not align with these findings, likely due to the difficulties with identification in these admixed sites, highlighting the intrinsic difficulty of defining subspecies boundaries. There was further substructure within *O. pannosa* subsp. *cardiophylla*, which is represented by several isolated and geographically distant collection localities. Our findings in Victoria are in agreement with the results of Smith, James and Ladiges [34] who also found that the collection site Rushworth (RUS) was genetically differentiated from the Brisbane Ranges (BR) and Anglesea (ANG) sites which could be explained by the large geographic distance between these localities (~150 km and ~220 km respectively).

Overall, the genetic structure results reflect the life history traits and disjunct range of *O. pannosa* as gene flow will be limited, resulting in greater genetic drift. Therefore, the strong differentiation in isolated collection localities is unsurprising. Similar patterns have been found in other threatened species [67,68]. For example, genomic analysis of the endangered Australian daisy (*Rutidosis leptorrhynchoides*) found strong genetic differentiation between geographically isolated collection localities [69].

4.2. Leaf Shape and Taxonomy

Our leaf shape analysis showed that the three main genetic clusters display relatively distinct leaf morphologies. Specimens measured from the *O. pannosa* subsp. *pannosa* genetic cluster had smaller, more oval shaped leaves compared to plants belonging to the *Olearia* (FR) and *O. pannosa* subsp. *cardiophylla* genetic groups, which we found to overlap more in size than shape. Previous taxonomic work on the species has classified subspecies by morphological traits and demonstrated clear differences between the leaf shapes of *O. pannosa* subsp. *pannosa* in South Australia and *O. pannosa* subsp. *cardiophylla* in Victoria [34], similar to what we find here. However, this previous analysis showed that *Olearia* (FR) samples (previously *O. pannosa* subsp. *cardiophylla* in South Australia) clustered with *O. pannosa* subsp. *cardiophylla* samples from Anglesea (ANG), Victoria. Here, we show that whilst the leaf shape of *Olearia* (FR) is more similar to *O. pannosa* subsp. *cardiophylla* than to *O. pannosa* subsp. *pannosa*, there are differences between them, with *Olearia* (FR) having more heart-shaped leaves. Geographic distance between *Olearia* (FR) and *O. pannosa* subsp. *cardiophylla* is extremely large (~315 km between Melrose (MEL) and the closest *O. pannosa* subsp. *cardiophylla* locality at Newland Head (NEW)). We also show that genetic divergence is high between *Olearia* (FR) and all other *O. pannosa* collection localities, suggesting they may be a distinct species, although reproductive isolation is yet to be tested.

We observed evidence for admixture among *O. pannosa* subsp. *pannosa* and *O. pannosa* subsp. *cardiophylla* on the Fleurieu Peninsula and in parts of the Southeast of SA, indicating that gene flow is possible between these groups. If Olearia (FR) and *O. pannosa* subsp. *pannosa* were not reproductively isolated, then the strong signal of genetic structure we observe between such geographically close collection localities inhabiting very similar environments is difficult to explain. The advance in sequencing technologies in recent years has led to numerous studies that have unearthed previously unrecognised 'cryptic species' [71–79]. These are genetically distinct taxa that are classified as the same species due to morphological similarity [80]. This species complex has undergone several taxonomic revisions since its first description by Hooker in 1851 (see Smith, James and Ladiges [34] and references therein), reflecting the large amount of phenotypic variation within the species. These results are also consistent with the ongoing difficulties with taxonomy and phylogeny in the *Olearia* genus [29–33]. Our findings strongly suggest a cryptic species complex in *O. pannosa*. For effective threatened species management plans, a firm understanding of

likely species or subspecies boundaries is essential, and our results highlight the importance of combining both genetic and morphological results methods for this approach.

4.3. Environment Likely Shapes Genotype, Phenotype, and Distribution

We conducted further analyses to identify the effect of the environment across the genotype, phenotype, and distribution of *O. pannosa* subsp. *pannosa.* These analyses identified maximum temperature as a key environmental variable that is a likely selective pressure on this group. Our analyses of genetic and leaf shape variation both identified maximum temperature as explaining a proportion of the measured variation. These findings are consistent with a number of studies that have found genetic signals [81–85] and morphological variation [86–90] associated with temperature being an agent of selection on plant species. Furthermore, the current distribution of *O. pannosa* subsp. *pannosa* appears to be closely related to climate. Our climate suitability models indicated specific constraints with respect to mean annual rainfall and summer maximum temperature gradients, with the species occurring within a transition zone between the more mesic and arid zones of the surrounding region [91]. Many collection localities of *O. pannosa* subsp. *pannosa* sit at the warm/arid thresholds of these key climatic gradients beyond which modelled probability of occurrence drops off markedly.

Taken together these results suggest that the current climatic trends towards higher summer maxima and increasing aridity are likely to have a detrimental impact on many collection localities. However, the few localities in cooler, wetter habitats may, in fact, benefit in the short term from climate change, leading to shifts in distribution and abundance. Combined, these results highlight the potential vulnerability of *O. pannosa* subsp. *pannosa* to the imminent effects of climate change and should be considered when planning conservation strategies for the species.

4.4. Conservation Importance and Implications for Each Genetic Group

4.4.1. *Olearia pannosa* (FR)

Olearia pannosa (FR) had low levels of genetic diversity across both collection localities and the collection locality MEL (Melrose) had the highest level of inbreeding across the whole dataset. We also found moderately high genetic differentiation between the two collection localities Dutchman's Stern (DUT) and Melrose (MEL). Concerningly, as there are only a few known stands of these individuals, all with small census estimates, this suggests an endangered species, or at least a subspecies which contains distinct genetic variation that will be lost if these they are not conserved. An immediate taxonomic review is essential to address these issues. We suggest a multi-level management strategy centred for this group. Firstly, we recommend in-situ threat abatement (i.e., protection from grazing pressure) and natural recruitment (i.e., creating disturbance to encourage recovery from the seedbank) [92]. We also recommend developing a strategy to aid the recovery of genetic diversity through a translocation plan which facilitates gene flow between the two isolated stands [92]. This could be developed through the ex-situ establishment of a managed seed production orchard to provide high quality seed [93]. Furthermore, further research could be informative, such as whole genome sequencing and breeding experiments, to ascertain (a) whether *Olearia* (FR) truly is a distinct, isolated species from *O. pannosa* and (b) what the functional genomic differences are between these divergent taxa.

4.4.2. *Olearia pannosa* subsp. *pannosa*

Within *O. pannosa* subsp. *pannosa*, there is evidence of some degree of historical gene flow across collection localities and low population substructure. This genetic group had higher levels of genetic diversity compared to *O. pannosa* subsp. *cardiophylla.* Observed heterozygosity was consistently lower than expected heterozygosity and was reflected in the significantly higher levels of historical inbreeding in *O. pannosa* subsp. *pannosa* than the other two genetic groups. Inbreeding has previously been flagged as a concern for *O. pannosa* [34] and historical selfing in this group may be one explanation for the results.

Due to the substructure that we found between isolated collection localities on the Eyre Peninsula (Coulta (COUL) Cummins (CM) and Cleeves (CLE)) and the rest of the group, we recommend that management efforts be focused here. Facilitating gene flow between these stands through translocations or increasing connectivity between may be advantageous.

As our results identify maximum temperature as potential agent of selection for this group, stands in more arid regions may experience detrimental effects from climate change. We recommend close monitoring, particularly of the more northern and arid localities. Apart from collection localities on the Eyre Peninsula, there is evidence of gene flow and higher levels of genetic diversity compared to other genetic groups. For these stands, we recommend a management approach which maintains connectivity and gene flow among localities. However, the potential future effect of aridification on this genetic group should be considered in future management strategies, and a climate-adjusted seed sourcing strategy [18] could be implemented (i.e., introducing seed from more arid areas to southern stands to bolster their adaptive capacity to arid conditions).

4.4.3. *Olearia pannosa* subsp. *cardiophylla*

We found that *O. pannosa* subsp. *cardiophylla* had the lowest levels of both historical inbreeding and genetic diversity, yet the highest levels of kinship. We also found that three of the four sites which had lowest genetic diversity and inbreeding in this group were the only collection sites that had to have highly related individuals removed across the whole dataset. There was considerable substructure in this genetic group, particularly between sites in Southeast SA and Victoria, and between sites within Victoria. It is likely that the large range disjunction of this group is driving genetic drift and genetic divergence. To develop an appropriate conservation management plan for *O. pannosa* subsp. *cardiophylla*, there are several knowledge gaps which still need to be addressed, particularly around the genetic divergence we observed between *O. pannosa* subsp. *cardiophylla* and the two other genetic groups, and the consequences that the high levels of substructure we found within this genetic group may have on mixing seed between localities (e.g., outbreeding depression). We suggest the establishment of a common garden trial which explores the effects on fitness and genetic diversity of genetic mixing between localities within this genetic group and amongst the three genetic groups. Until this information is known, in a scenario with high structure but low genetic diversity and inbreeding, Ottewell, Bickerton, Byrne, and Lowe [88] suggests a cautious management approach to minimise the risk of outbreeding depression with a focus on in-situ recovery of genetic diversity through the seed bank and a composite translocation strategy [19].

5. Conclusions

In summary, combining range-wide genomic and leaf shape analysis of the threatened Australian daisy *Olearia pannosa* revealed a likely cryptic species (*Olearia* (FR)) and identified the need to redefine the geographic intergrade between existing subspecific boundaries (*O. pannosa* subsp. *cardiophylla* and *O. pannosa* subsp. pannosa). Levels of genetic diversity and inbreeding differed between the three genetic groups, suggesting each requires a separate management strategy which we define below.

- *Olearia* (FR)–We suggest a management strategy centred around in-situ threat abatement and disturbance to encourage seedbank recovery, the recovery of genetic diversity through a translocation plan to encourage gene flow between the two isolated stands, and the development of a seed production orchard to supply the seed.
- *O. pannosa* subsp. *pannosa*—The main priority is to facilitate gene flow between existing stands or to increase connectivity between them, especially in an arid-to-mesic direction since it appears that maximum temperature is an important agent of selection.
- *O. pannosa* subsp. *cardiophylla*—It is likely that the large range disjunction of this group is driving genetic divergence. To develop an appropriate conservation management plan, several knowledge gaps still need to be addressed, particularly to assess the

potential for outbreeding depression. Until then, we recommend priority is given to in-situ recovery of genetic diversity.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/life11060553/s1, Supplementary File S1—The variant call format (VCF) file will be made available as supplementary (Olearia_pannosa.vcf). Reads and mapping files will be archived at the NCBI SRA (accession number TBA). Supplementary File S2—the covariates from the redundancy analysis (genetic_RDA_covariates.xlsx).

Author Contributions: Conceptualization, C.B., D.C.B., M.F.B. and A.J.L.; Data curation, C.B., M.J.C. and M.F.B.; Formal analysis, C.B., M.J.C., N.R.F., G.R.G., A.R.G.M. and A.J.L.; Funding acquisition, C.B., M.F.B. and A.J.L.; Methodology, C.B., M.J.C., D.C.B., M.F.B., G.R.G., A.R.G.M. and A.J.L.; Project administration, C.B., M.F.B., D.C.B. and A.J.L.; Resources, C.B., M.F.B. and A.J.L.; Supervision, M.J.C., M.F.B. and A.J.L.; Visualization, C.B., M.J.C., M.F.B. and A.J.L.; Writing—Original draft, C.B.; Writing—review and editing, C.B., M.J.C., D.C.B., M.F.B., N.R.F., G.R.G., A.R.G.M. and A.J.L.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded as part of the 'Securing the Future of the Silver Daisy Project' (SSDP), which was managed by Trees for Life (South Australia) and funded through the Australian Government's 'Threatened Species Recovery Fund' (2016–2017). This project was also supported by The Holsworth Wildlife Research Endowment & The Ecological Society of Australia.

Data Availability Statement: Reads and mapping files will be archived at the NCBI SRA (accession number TBA).

Acknowledgments: The authors wish to thank Amelia Hurren, Rohan Cleeves, Randall Bates, and Vicki-Jo Rusell from Trees for Life (SA), Rani Hunt from the Department of Environment, Land, Water and Planning (Victoria) and Chris Lindorff from Trust for Nature for providing their assistance and expertise throughout this and for assisting us with sample collection. We'd like to thank Peter Lang and Helen Vonow from the Sate Herbarium of South Australia for providing their expertise in identification of vouchers specimens and advise on taxonomy. Thanks to Dan Duval at the South Australian Seed Conservation Centre for sharing advice and expertise. Thanks also go to Sam Turner-Blyth and Kiri Marker for their assistance with drafting and editing the manuscript. Finally, a special thanks to Jeremy Austin for his assistance behind the scenes.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

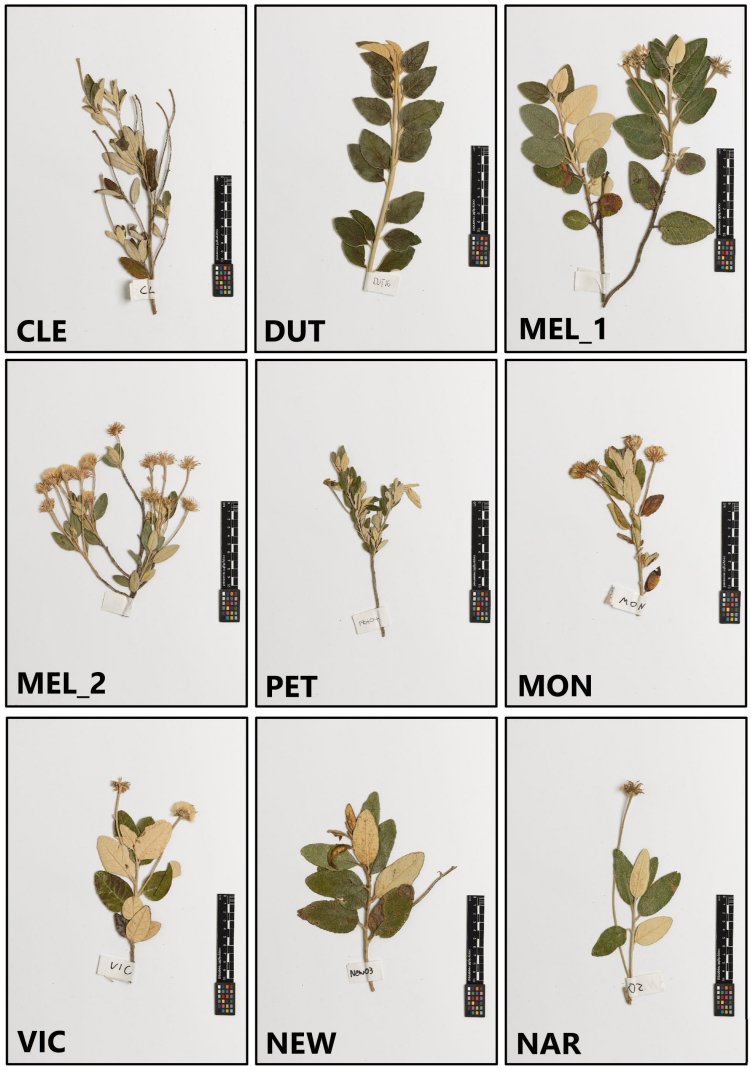


Figure A1. Examples of vouchers used in the leaf shape analysis. Vouchers selected represent the range of variation found across collection localities assigned to each genetic group. *Olearia* (FR) examples are from Melrose (MEL_1) and Dutchmans Stern (DUT). *O. pannosa* subsp. *pannosa* is represented by the collection localities Cleeves (CLE), Melrose (MEL_2), and Peterborough (PET). Victor Harbor (VIC) and Newland Head (NEW) are both localities which have admixture between *O. pannosa* subsp. *pannosa* and *O. pannosa* subsp. *cardiophylla*. The *O. pannosa* subsp. *cardiophylla* example is from Naracoorte (NAR).

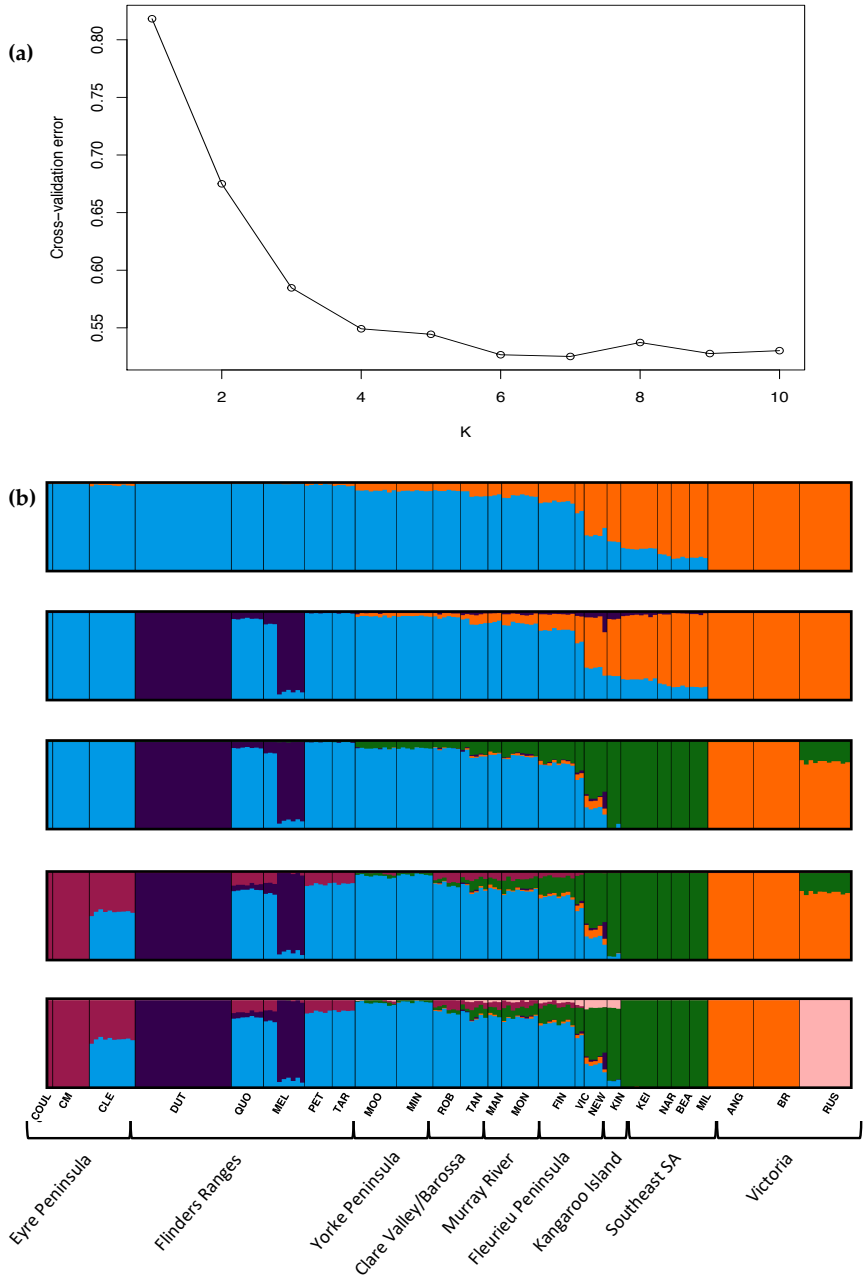


Figure A2. (**a**) ADMIXTURE error results—the cv errors are plotted for K = 1–10, with the lowest cv error indicating greatest support at K = 6. (**b**) Bar plots represent individual genetic cluster assignment from ADMIXTURE results from K = 2 to K = 6 (after which the assignment becomes messy and incomprehensible).

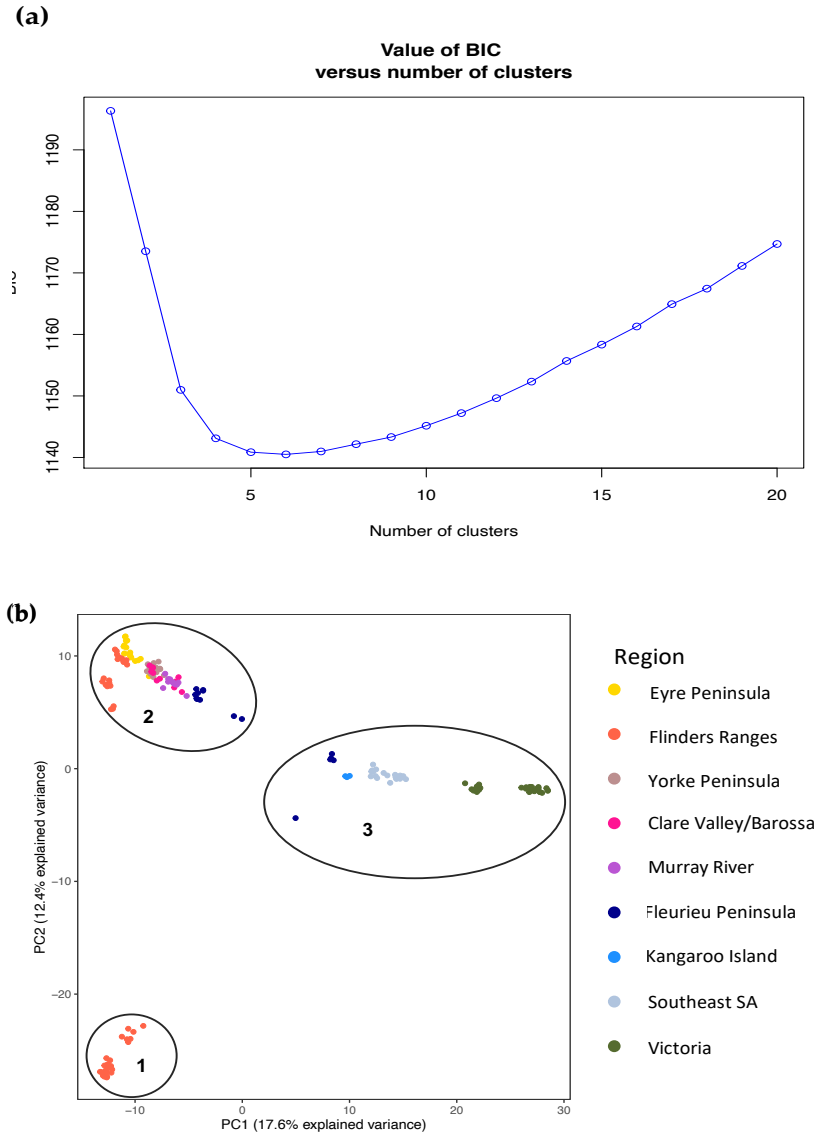


Figure A3. (**a**) Bayesian Information Criterion (BIC) values for each number of clusters, with the lowest BIC indicating support for five or six clusters; (**b**) Principal Component Analysis showing population genetic structure across all samples. Colours represent regions and ellipses identify the genetic clusters (1–3) identified in the DAPC analysis for K = 3.

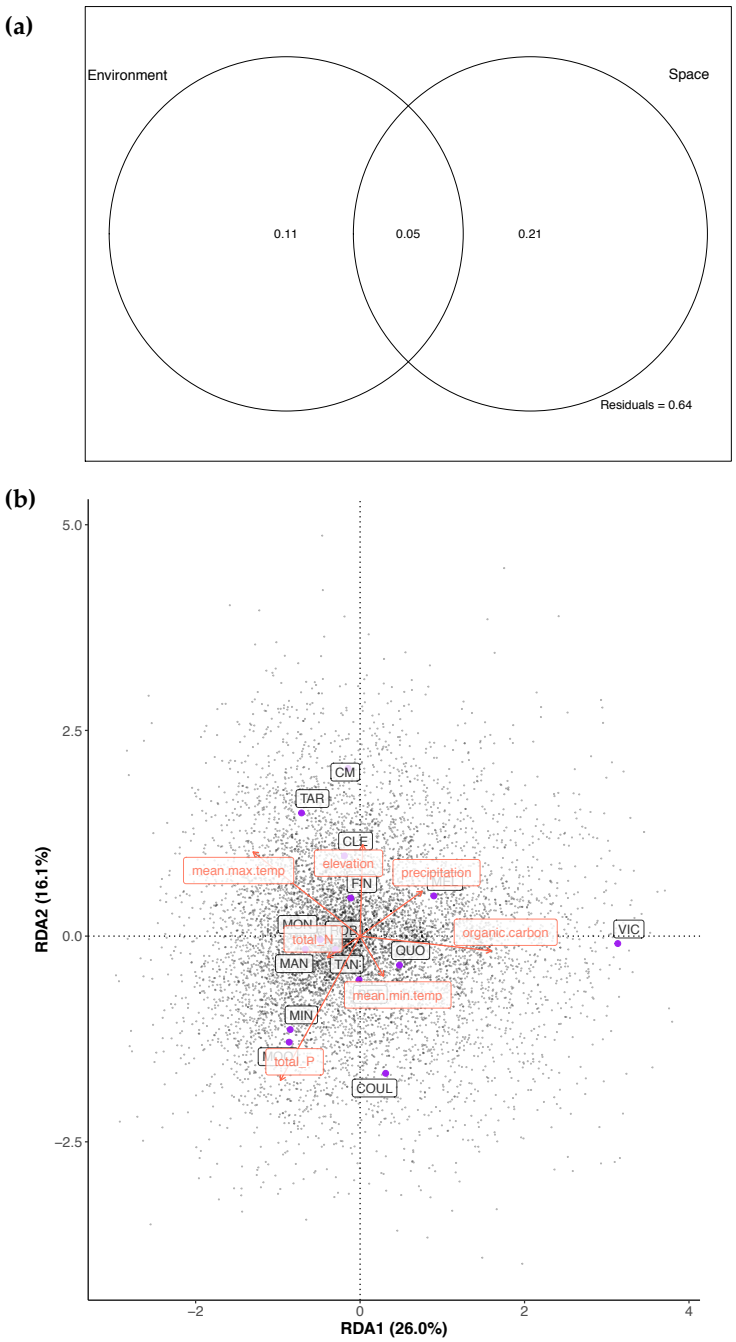


Figure A4. (**a**) Venn diagram showing results from variation partitioning, with environmental variables and spatial distribution used as explanatory variables, showing that 11% of the genetic variance is uniquely explained by environment. (**b**) Triplot for the first two axis of a partial redundancy analysis with environment as explanatory variables conditioned on spatial distribution.

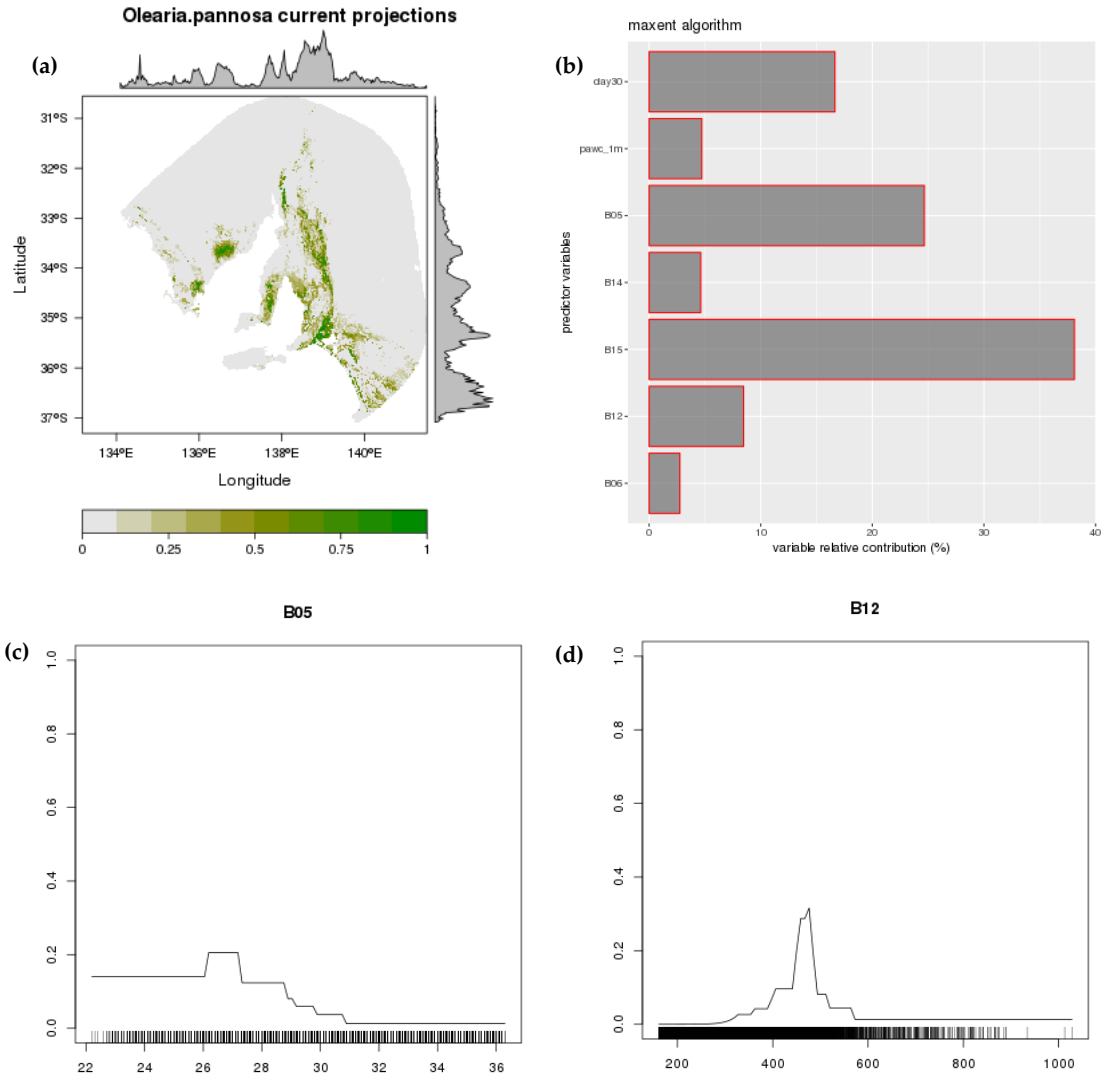


Figure A5. Projected distribution of *O. pannosa* subsp. *pannosa* (MaxEnt). Current distribution (**a**); variable relative contribution (**b**); two example response curves (mean maximum temperature of the warmest month; (**c**) and mean annual rainfall; (**d**)).

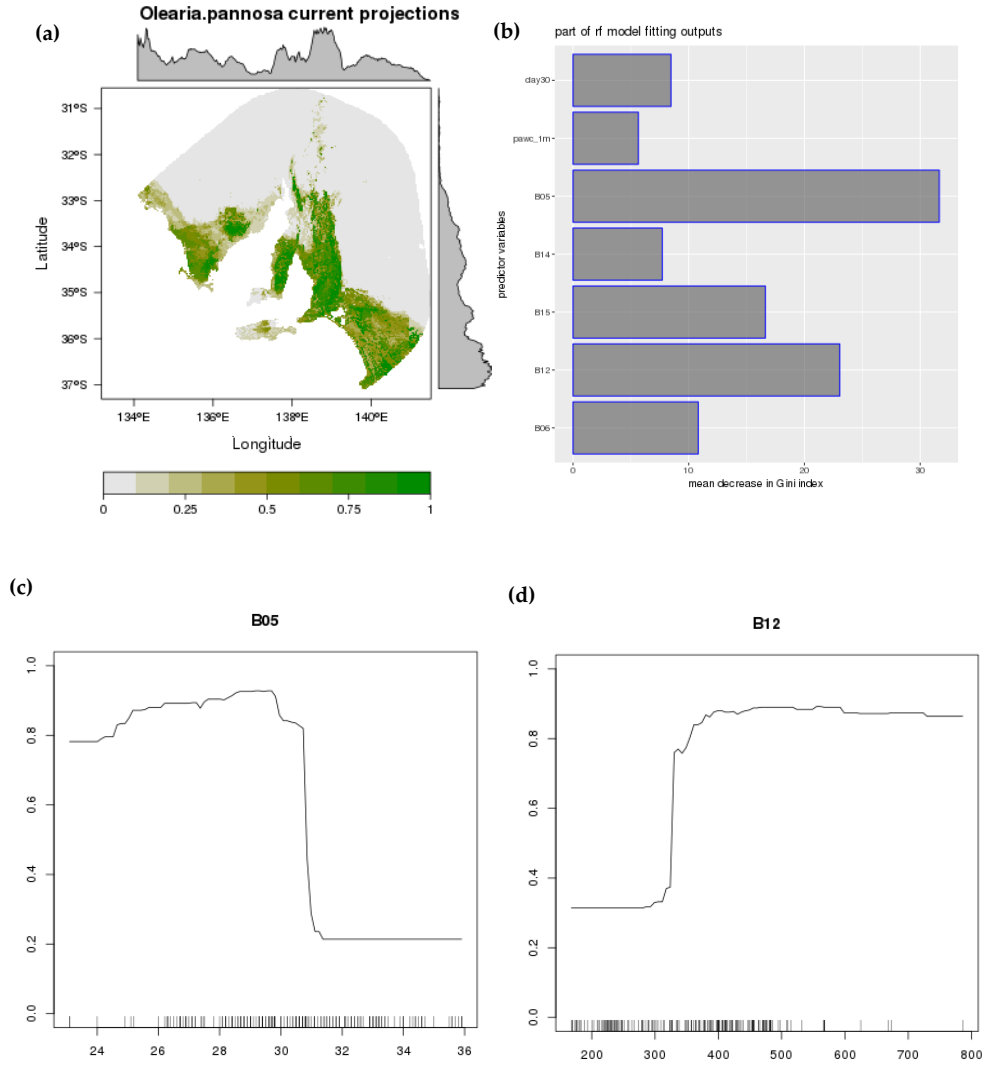


Figure A6. Projected distribution of *O. pannosa* subsp. *pannosa* (Random Forest). Current distribution (**a**); variable importance (**b**); two example response curves (mean maximum temperature of the warmest month; (**c**) and mean annual rainfall; (**d**)).

Table A1. Summary of region, collection locality, collection ID, collection site coordinates, locality voucher identification (provided by the State Herbarium of South Australia, locality genetic group assignment, and the 2018 census estimate for each collection locality

Region	Collection Locality	Collection ID	Latitude	Longitude	Subspecies (SA Herbarium ID)	Genetic Group	2018 Census Estimate
Eyre Peninsula	Coulta	COUL	−34.38945	135.4347	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	1
	Cummins	CM	−34.32097	135.9643	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	440
	Cleeve	CLE	−33.6998	136.5024	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	490
Flinders Ranges	Dutchman's Stern	DUT	−32.30784	137.9728	*O. pannosa* subsp. *cardiophylla*	*O. pannosa* (FR)	400
	Quorn	QUO	−32.49042	138.0478	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	6300
	Melrose	MEL	−32.69477	138.1048	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	231
	Melrose	MEL	−32.71896	138.0935	*O. pannosa* subsp. *cardiophylla*	*O. pannosa* (FR)	345
	Peterborough	PET	−32.78896	138.8835	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	60
	Tarcowie	TAR	−34.65343	138.8208	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	29
Yorke Peninsula	Moonta	MOO	−34.54376	137.7825	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	370
	Minlaton	MIN	−34.6748	137.7199	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	385
Clare Valley/Barossa	Robertstown	ROB	−33.85449	139.0413	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	1523
	Tanunda	TAN	−34.38884	138.8028	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	368
Murray River	Mannum	MAN	−34.8747	139.2146	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	22
	Monarto	MON	−35.22497	139.0047	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	301
Fleurieu Peninsula	Finniss	FIN	−35.31793	138.8836	*O. pannosa* subsp. *pannosa*	*O. pannosa* subsp. *pannosa*	513
	Victor Harbor	VIC	−35.47631	138.7684	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *pannosa*	115
	Newland Head	NEW	−35.61022	138.5018	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	31
Kangaroo Island	Kingscote	KIN	−35.57406	137.576	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	12
Southeast SA	Keith	KEI	−36.14328	140.4094	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	514
	Naracoorte	NAR	−37.11145	140.5361	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	80
	Beachport	BEA	−37.28313	139.9409	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	14
	Millicent	MIL	−37.68687	140.5147	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	18
Victoria	Anglesea	ANG	−38.37622	144.245	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	Unknown
	Brisbane Ranges	BR	−37.90723	144.232	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	Unknown
	Rushworth	RUS	−36.62514	145.1126	*O. pannosa* cf. subsp. *cardiophylla*	*O. pannosa* subsp. *cardiophylla*	Unknown

Table A2. Site ID, genetic group, highly related individuals removed from collection locality, sample size after filtering (n), number of variant sites after filtering for a minor allele count of 1, observed (HO) and expected (HE) heterozygosity (both vary from 0 to 1, with 0 indicating no diversity), inbreeding coefficient (*F*) (varies from 0 to 1 and values of 0 indicate no evidence of inbreeding), and Kinship (0.5 indicates samples are genetically identical. The lower the kinship coefficient, the lower the level of kinship between individuals).

Collection ID	Genetic Group	Highly Related Individuals Removed	n	Variant Sites	H_O	H_E	F	Kinship
COUL	*O. pannosa* subsp. *pannosa*	-	1	1379	—	—	—	—
CM	*O. pannosa* subsp. *pannosa*	-	8	5635	0.174 0.173 to 0.175)	0.221 (0.220 to 0.222)	0.213 (0.208 to 0.218)	−0.239 (−0.274 to −0.211)
CLE	*O. pannosa* subsp. *pannosa*	-	10	7479	0.218 (0.217 to 0.219)	0.285 (0.284 to 0.286)	0.234 (0.231 to 0.237)	−0.308 (−0.328 to −0.280)
DUT	*O. pannosa* (FR)	-	21	4612	0.17 (0.169 to 0.171)	0.224 (0.223 to 0.225)	0.241 (0.238 to 0.244)	−0.216 (−0.231 to −0.202)
QUO	*O. pannosa* subsp. *pannosa*	-	7	6457	0.19 (0.189 to 0.191)	0.254 (0.253 to 0.255)	0.253 (0.249 to 0.257)	−0.324 (−0.347 to −0.303)
MEL_2	*O. pannosa* subsp. *pannosa*	-	3	4931	0.194 (0.191 to 0.197)	0.268 (0.265 to 0.271)	0.276 (0.266 to 0.286)	−0.394 (−0.536 to −0.154)
MEL_1	*O. pannosa* (FR)	-	6	3968	0.164 (0.162 to 0.166)	0.243 (0.241 to 0.245)	0.325 (0.318 to 0.332)	−0.388 (−0.478 to −0.315)
PET	*O. pannosa* subsp. *pannosa*	-	6	6666	0.217 (0.215 to 0.219)	0.275 (0.273 to 0.277)	0.211 (0.206 to 0.216)	−0.242 (−0.265 to −0.212)
TAR	*O. pannosa* subsp. *pannosa*	-	5	6468	0.22 (0.218 to 0.222)	0.281 (0.279 to 0.283)	0.218 (0.212 to 0.224)	−0.264 (−0.301 to −0.204)
MOO	*O. pannosa* subsp. *pannosa*	-	9	7613	0.219 (0.218 to 0.220)	0.293 (0.292 to 0.294)	0.250 (0.247 to 0.253)	−0.326 (−0.351 to −0.307)
MIN	*O. pannosa* subsp. *pannosa*	-	8	7254	0.217 (0.216 to 0.218)	0.286 (0.285 to 0.287)	0.241 (0.237 to 0.245)	−0.312 (−0.343 to −0.284)
ROB	*O. pannosa* subsp. *pannosa*	-	6	7337	0.231 (0.229 to 0.233)	0.304 (0.302 to 0.306)	0.241 (0.236 to 0.246)	−0.307 (−0.326 to −0.292)
TAN	*O. pannosa* subsp. *pannosa*	-	6	7645	0.231 (0.229 to 0.233)	0.317 (0.315 to 0.319)	0.270 (0.265 to 0.275)	−0.370 (−0.406 to −0.331)
MAN	*O. pannosa* subsp. *pannosa*	-	3	5571	0.228 (0.225 to 0.231)	0.300 (0.297 to 0.303)	0.240 (0.231 to 0.249)	−0.298 (−0.308 to −0.290)
MON	*O. pannosa* subsp. *pannosa*	-	8	8299	0.236 (0.235 to 0.237)	0.322 (0.321 to 0.323)	0.268 (0.265 to 0.271)	−0.394 (−0.436 to −0.364)
FIN	*O. pannosa* subsp. *pannosa*	-	8	8224	0.245 (0.244 to 0.246)	0.319 (0.318 to 0.320)	0.232 (0.229 to 0.235)	−0.295 (−0.320 to −0.271)
VIC	*O. pannosa* subsp. *cardiophylla*	-	2	4694	0.232 (0.228 to 0.236)	0.330 (0.326 to 0.334)	0.298 (0.286 to 0.310)	−0.476 (−0.476 to −0.476)
NEW	*O. pannosa* subsp. *cardiophylla*	-	5	6485	0.224 (0.222 to 0.226)	0.315 (0.313 to 0.317)	0.288 (0.282 to 0.294)	−0.499 (−0.692 to −0.338)
KIN	*O. pannosa* subsp. *cardiophylla*	1	2	1393	0.096 (0.093 to 0.099)	0.109 (0.106 to 0.112)	0.117 (0.093 to 0.141)	−0.165 (−0.165 to −0.165)
KEI	*O. pannosa* subsp. *cardiophylla*	-	8	5436	0.172 (0.171 to 0.173)	0.239 (0.238 to 0.240)	0.279 (0.274 to 0.284)	−0.350 (−0.387 to −0.323)
NAR	*O. pannosa* subsp. *cardiophylla*	-	3	7200	0.159 (0.156 to 0.162)	0.184 (0.181 to 0.187)	0.137 (0.125 to 0.149)	−0.156 (−0.194 to −0.094)
BEA	*O. pannosa* subsp. *cardiophylla*	-	4	3552	0.157 (0.154 to 0.160)	0.190 (0.187 to 0.193)	0.174 (0.164 to 0.184)	−0.213 (−0.266 to −0.168)
MIL	*O. pannosa* subsp. *cardiophylla*	-	4	3329	0.153 (0.150 to 0.156)	0.180 (0.177 to 0.183)	0.148 (0.138 to 0.158)	−0.156 (−0.179 to −0.137)
ANG	*O. pannosa* subsp. *cardiophylla*	2	8	2525	0.107 (0.106 to 0.108)	0.113 (0.112 to 0.114)	0.055 (0.049 to 0.061)	−0.062 (−0.091 to −0.029)
BR	*O. pannosa* subsp. *cardiophylla*	3	7	3081	0.127 (0.126 to 0.128)	0.138 (0.137 to 0.139)	0.078 (0.072 to 0.084)	−0.077 (−0.090 to −0.059)
RUS	*O. pannosa* subsp. *cardiophylla*	-	11	3875	0.145 (0.144 to 0.146)	0.166 (0.165 to 0.167)	0.127 (0.123 to 0.131)	−0.129 (−0.159 to −0.089)

95% confidence intervals in parentheses.

Table A3. Pairwise genetic differentiation (measured by FST) between collection localities (values vary from 0–1 and values of 0 indicate no evidence of differentiation)

	COUL	CM	CLE	DUT	QUO	MEL_2	MEL_1	PET	TAR	MOO	MIN	ROB	TAN	MAN	MON	FIN	VIC	NEW	KIN	KEI	NAR	BEA	MIL	ANG	BR	RUS
COUL	0.00																									
CM	0.05	0.00																								
CLE	0.07	0.13	0.00																							
DUT	0.63	0.58	0.49	0.00																						
QUO	0.20	0.25	0.13	0.51	0.00																					
MEL_2	0.19	0.27	0.13	0.52	0.07	0.00																				
MEL_1	0.59	0.53	0.42	0.23	0.45	0.45	0.00																			
PET	0.17	0.22	0.10	0.52	0.09	0.10	0.45	0.00																		
TAR	0.16	0.21	0.09	0.52	0.09	0.09	0.45	0.06	0.00																	
MOO	0.15	0.22	0.11	0.49	0.15	0.15	0.41	0.12	0.11	0.00																
MIN	0.17	0.23	0.12	0.51	0.17	0.16	0.43	0.13	0.12	0.00	0.00															
ROB	0.11	0.20	0.07	0.50	0.12	0.11	0.41	0.08	0.06	0.07	0.08	0.00														
TAN	0.09	0.20	0.07	0.48	0.12	0.10	0.39	0.08	0.06	0.06	0.07	0.01	0.00													
MAN	0.14	0.25	0.10	0.54	0.16	0.13	0.46	0.11	0.09	0.09	0.10	0.05	0.03	0.00												
MON	0.08	0.19	0.07	0.46	0.11	0.09	0.37	0.07	0.05	0.06	0.07	0.01	0.00	0.02	0.00											
FIN	0.12	0.21	0.09	0.46	0.13	0.12	0.37	0.10	0.08	0.08	0.09	0.04	0.02	0.04	0.02	0.00										
VIC	0.11	0.30	0.15	0.55	0.21	0.17	0.46	0.16	0.14	0.12	0.13	0.08	0.04	0.07	0.03	0.03	0.00									
NEW	0.29	0.36	0.26	0.52	0.31	0.29	0.43	0.28	0.26	0.22	0.23	0.20	0.16	0.19	0.15	0.12	0.09	0.00								
KIN	0.72	0.52	0.39	0.66	0.46	0.49	0.64	0.43	0.42	0.36	0.38	0.36	0.31	0.41	0.29	0.27	0.36	0.22	0.00							
KEI	0.46	0.46	0.36	0.59	0.41	0.42	0.53	0.39	0.38	0.33	0.34	0.33	0.29	0.34	0.27	0.25	0.27	0.17	0.36	0.00						
NAR	0.58	0.50	0.38	0.64	0.44	0.46	0.60	0.42	0.41	0.35	0.36	0.35	0.31	0.38	0.28	0.27	0.32	0.20	0.51	0.17	0.00					
BEA	0.57	0.50	0.39	0.63	0.45	0.48	0.59	0.43	0.42	0.36	0.37	0.36	0.33	0.40	0.30	0.28	0.34	0.20	0.47	0.16	0.12	0.00				
MIL	0.59	0.51	0.40	0.64	0.46	0.49	0.60	0.44	0.43	0.37	0.38	0.37	0.33	0.41	0.31	0.28	0.35	0.22	0.50	0.18	0.16	0.12	0.00			
ANG	0.75	0.63	0.53	0.71	0.59	0.65	0.71	0.58	0.59	0.51	0.53	0.53	0.50	0.60	0.47	0.46	0.58	0.45	0.68	0.49	0.59	0.56	0.58	0.00		
BR	0.70	0.60	0.50	0.69	0.56	0.61	0.68	0.55	0.55	0.49	0.50	0.50	0.47	0.56	0.44	0.43	0.53	0.41	0.63	0.45	0.55	0.52	0.54	0.20	0.00	
RUS	0.64	0.57	0.47	0.66	0.53	0.56	0.64	0.52	0.52	0.46	0.47	0.47	0.44	0.51	0.41	0.40	0.47	0.36	0.55	0.42	0.49	0.47	0.48	0.46	0.42	0.00

Table A4. Kinship values between highly related individuals removed from the dataset (0.5 indicates samples are genetically identical. The lower the kinship coefficient, the lower the level of kinship between individuals).

Collection Locality	Collection ID Individual 1	Collection Locality	Collection ID Individual 2	Kinship	Relation
KIN	KIN_03	KIN	KIN_02	0.497758	clone
BR	BR_04	BR	BR_02	0.265698	first-degree
ANG	ANG_02	ANG	ANG_01	0.254376	first-degree
ANG	ANG_05	ANG	ANG_04	0.252064	first-degree
BR	BR_09	BR	BR_05	0.241582	first-degree
RUS	RUS_06	RUS	RUS_01	0.212617	first-degree
ANG	ANG_10	ANG	ANG_09	0.169238	second-degree
RUS	RUS_11	RUS	RUS_04	0.120962	second-degree
ANG	ANG_10	ANG	ANG_07	0.107917	second-degree
ANG	ANG_09	ANG	ANG_07	0.103653	second-degree
BR	BR_02	BR	BR_01	0.103425	second-degree
ANG	ANG_10	ANG	ANG_06	0.102982	second-degree
BR	BR_04	BR	BR_01	0.101133	second-degree

Table A5. Results of the Tukey pairwise comparison to detect differences in observed heterozygosity (H_O), expected heterozygosity (H_E) and the inbreeding coefficient (*F*) between the three genetic groups *O. pannosa* subsp. *pannosa, Olearia* (FR) and *O. pannosa* subsp. *Cardiophylla.* Diff is the difference between means of the two groups, lwr and upr are the ends of the 95% confidence interval and p adj is the *p*-value after adjusting for the comparisons.

	Group Comparison	diff	lwr	upr	*p* adj
H_O	*O. pannosa* subsp. *cardiophylla—Olearia* (FR)	−0.02	−0.07	0.04	0.68
	O. pannosa subsp. *pannosa—Olearia* (FR)	0.05	0	0.1	0.05
	O. pannosa subsp. *pannosa—O. pannosa* subsp. *cardiophylla*	0.07	0.04	0.1	<0.001
H_E	O. *pannosa* subsp. *cardiophylla—Olearia* (FR)	−0.05	−0.14	0.04	0.32
	O. pannosa subsp. *pannosa—Olearia* (FR)	0.06	−0.03	0.14	0.25
	O. pannosa subsp. *pannosa—O. pannosa* subsp. *cardiophylla*	0.11	0.06	0.16	<0.001
F	*O. pannosa* subsp. *cardiophylla—Olearia* (FR)	−0.13	−0.23	−0.02	0.02
	O. pannosa subsp. *pannosa—Olearia* (FR)	−0.04	−0.14	0.07	0.64
	O. *pannosa* subsp. *pannosa—O. pannosa* subsp. *cardiophylla*	0.09	0.03	0.15	<0.001
Kinship	*O. pannosa* subsp. *cardiophylla—Olearia* (FR)	0.10	−0.10	0.30	0.42
	O. *pannosa* subsp. *pannosa—Olearia* (FR)	−0.02	−0.22	0.17	0.95
	O. pannosa subsp. *pannosa—O. pannosa* subsp. *cardiophylla*	−0.12	−0.23	−0.02	0.02

Table A6. Results of the one-way analysis of variance (ANOVA) conducted to detect differences in leaf shape between the three genetic groups *O. pannosa* subsp. *pannosa, Olearia* (FR), and *O. pannosa* subsp. *cardiophylla.* We have compared the differences between PC1 (width, length, and leaf surface area), and PC2 (ovality and location of widest point of the leaf).

	Group Comparison	diff	lwr	upr	*p* adj
PC1	*O. pannosa* subsp. *cardiophylla—Olearia* (FR)	0.68	−0.18	1.54	0.15
	O. pannosa subsp. *pannosa—Olearia* (FR)	2.45	1.74	1.74	<0.001
	O. pannosa subsp. *pannosa—O. pannosa* subsp. *cardiophylla*	3.13	2.42	3.84	<0.001
PC2	*O. pannosa* subsp. *cardiophylla—Olearia* (FR)	1.94	1.33	2.56	<0.001
	O. pannosa subsp. *pannosa—Olearia* (FR)	0.11	−0.39	0.62	0.86
	O. pannosa subsp. *pannosa—O. pannosa* subsp. *cardiophylla*	−1.83	−2.33	−1.32	<0.001

References

1. Breed, M.F.; Stead, M.G.; Ottewell, K.M.; Gardner, M.G.; Lowe, A.J. Which provenance and where? Seed sourcing strategies for revegetation in a changing environment. *Conserv. Genet.* **2013**, *14*, 1–10. [CrossRef]
2. Breed, M.F.; Gardner, M.G.; Ottewell, K.; Navarro, C.; Lowe, A.J. Shifts in reproductive assurance strategies and inbreeding costs associated with habitat fragmentation in Central American mahogany. *Ecol. Lett.* **2012**, *15*, 444–452. [CrossRef] [PubMed]

3. Lowe, A.J.; Cavers, S.; Boshier, D.; Breed, M.F.; Hollingsworth, P.M. The resilience of forest fragmentation genetics—No longer a paradox—We were just looking in the wrong place. *Heredity* **2015**, *115*, 97–99. [CrossRef]
4. Eckert, C.G.; Kalisz, S.; Geber, M.A.; Sargent, R.; Elle, E.; Cheptou, P.-O.; Goodwillie, C.; Johnston, M.O.; Kelly, J.K.; Moeller, D.A.; et al. Plant mating systems in a changing world. *Trends Ecol. Evol.* **2010**, *25*, 35–43. [CrossRef] [PubMed]
5. Lowe, A.J.; Boshier, D.; Ward, M.; Bacles, C.F.E.; Navarro, C. Genetic resource impacts of habitat loss and degradation; Reconciling empirical evidence and predicted theory for neotropical trees. *Heredity* **2005**, *95*, 255–273. [CrossRef]
6. Aitken, S.N.; Yeaman, S.; Holliday, J.A.; Wang, T.; Curtis-McLane, S. Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evol. Appl.* **2008**, *1*, 95–111. [CrossRef] [PubMed]
7. Christmas, M.J.; Breed, M.F.; Lowe, A.J. Constraints to and conservation implications for climate change adaptation in plants. *Conserv. Genet.* **2015**, *17*, 305–320. [CrossRef]
8. Pachauri, R.K.; Allen, M.R.; Barros, V.R.; Broome, J.; Cramer, W.; Christ, R.; Church, J.A.; Clarke, L.; Dahe, Q.; Dasgupta, P. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; IPCC: Geneva, Switzerland, 2014.
9. Cook, J.; Nuccitelli, D.; Green, S.A.; Richardson, M.; Winkler, B.; Painting, R.; Way, R.; Jacobs, P.; Skuce, A. Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environ. Res. Lett.* **2013**, *8*, 024024. [CrossRef]
10. Parmesan, C. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **2006**, *37*, 637–669. [CrossRef]
11. Jump, A.; Peñuelas, J. Running to stand still: Adaptation and the response of plants to rapid climate change. *Ecol. Lett.* **2005**, *8*, 1010–1020. [CrossRef]
12. Bongaarts, J. *IPBES, Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*; Wiley Online Library: Hoboken, NJ, USA, 2019.
13. Corlett, R.T.; Westcott, D.A. Will plant movements keep up with climate change? *Trends Ecol. Evol.* **2013**, *28*, 482–488. [CrossRef] [PubMed]
14. Young, A.; Boyle, T.; Brown, T. The population genetic consequences of habitat fragmentation for plants. *Trends Ecol. Evol.* **1996**, *11*, 413–418. [CrossRef]
15. Whiteley, A.R.; Fitzpatrick, S.W.; Funk, W.C.; Tallmon, D.A. Genetic rescue to the rescue. *Trends Ecol. Evol.* **2015**, *30*, 42–49. [CrossRef]
16. Aitken, S.N.; Whitlock, M.C. Assisted gene flow to facilitate local adaptation to climate change. *Annu. Rev. Ecol. Evol. Syst.* **2013**, *44*, 367. [CrossRef]
17. Gellie, N.J.C.; Breed, M.F.; Thurgate, N.; Kennedy, S.A.; Lowe, A.J. Local maladaptation in a foundation tree species: Implications for restoration. *Biol. Conserv.* **2016**, *203*, 226–232. [CrossRef]
18. Prober, S.M.; Byrne, M.; McLean, E.H.; Steane, D.A.; Potts, B.M.; Vaillancourt, R.E.; Stock, W.D. Climate-adjusted provenancing: A strategy for climate-resilient ecological restoration. *Front. Ecol. Evol.* **2015**, *3*, 65. [CrossRef]
19. Broadhurst, L.M.; Lowe, A.; Coates, D.J.; Cunningham, S.A.; McDonald, M.; Vesk, P.A.; Yates, C. Seed supply for broadscale restoration: Maximizing evolutionary potential. *Evol. Appl.* **2008**, *1*, 587–597. [CrossRef]
20. Ramalho, C.E.; Byrne, M.; Yates, C.J. A Climate-Oriented Approach to Support Decision-Making for Seed Provenance in Ecological Restoration. *Front. Ecol. Evol.* **2017**, *5*, 95. [CrossRef]
21. Bucharova, A.; Bossdorf, O.; Hölzel, N.; Kollmann, J.; Prasse, R.; Durka, W. Mix and match: Regional admixture provenancing strikes a balance among different seed-sourcing strategies for ecological restoration. *Conserv. Genet.* **2019**, *20*, 7–17. [CrossRef]
22. Breed, M.F.; Harrison, P.A.; Blyth, C.; Byrne, M.; Gaget, V.; Gellie, N.J.; Groom, S.V.; Hodgson, R.; Mills, J.G.; Prowse, T.A. The potential of genomics for restoring ecosystems and biodiversity. *Nat. Rev. Genet.* **2019**, *20*, 615–628. [CrossRef]
23. Breed, M.; Harrison, P.; Bischoff, A.; Durruty, P.; Gellie, N.; Gonzales, E.; Havens, K.; Kilkenny, F.; Krauss, S.; Lowe, A. Priority actions to improve provenance decision making. *Bioscience* **2018**, *68*, 510–516. [CrossRef]
24. Ledig, F.T.; Rehfeldt, G.E.; Jaquish, B. Projections of suitable habitat under climate change scenarios: Implications for transboundary assisted colonization. *Am. J. Bot.* **2012**, *99*, 1217–1230. [CrossRef] [PubMed]
25. Wang, C.; Liu, C.; Wan, J.; Zhang, Z. Climate change may threaten habitat suitability of threatened plant species within Chinese nature reserves. *PeerJ* **2016**, *4*, e2091. [CrossRef]
26. Vitt, P.; Havens, K.; Kramer, A.T.; Sollenberger, D.; Yates, E. Assisted migration of plants: Changes in latitudes, changes in attitudes. *Biol. Conserv.* **2010**, *143*, 18–27. [CrossRef]
27. Vitt, P.; Havens, K.; Hoegh-Guldberg, O. Assisted migration: Part of an integrated conservation strategy. *Trends Ecol. Evol.* **2009**, *24*, 473–474. [CrossRef] [PubMed]
28. Commander, L.E.; Coates, D.; Broadhurst, L.; Offord, C.A.; Makinson, R.O.; Matthes, M. *Guidelines for the Translocation of Threatened Plants in Australia*; Australian Network for Plant Conservation Incorporated: Canberra, Australia, 2018.
29. Lyddiard, D.; Ozga, T.; Greatrex, B.W. Diterpene and flavonoid constituents of the newly identified Australian species Olearia fulgens. *Phytochem. Lett.* **2020**, *39*, 19–24. [CrossRef]
30. Cross, E.W.; Quinn, C.J.; Wagstaff, S.J. Molecular evidence for the polyphyly of Olearia (Astereae: Asteraceae). *Plant Syst. Evol.* **2002**, *235*, 99–120. [CrossRef]
31. Messina, A.; Walsh, N.G.; Hoebee, S.E.; Green, P.T. A morphological assessment of the Olearia phlogopappa complex (Asteraceae: Astereae). *Aust. Syst. Bot.* **2013**, *26*, 31–80. [CrossRef]

32. Messina, A.; Walsh, N.G.; Hoebee, S.E.; Green, P.T. A revision of Olearia section Asterotriche (Asteraceae: Astereae). *Aust. Syst. Bot.* **2014**, *27*, 199–240. [CrossRef]
33. Walsh, N.G. Notes on Olearia (Asteraceae: Astereae) in south-east Australia: O. tenuifolia, O. adenophora and description of a new species endemic to eastern Victoria. *Muelleria* **2014**, *32*, 34–38.
34. Smith, Z.; James, E.A.; Ladiges, P.Y. Morphological and genetic variation in the rare daisy Olearia pannosa subsp. cardiophylla (Asteraceae). *Muelleria* **2004**, *20*, 33–48.
35. Hurren, A.; Breed, M.; Blyth, C.; Mason, A.; Cleeves, R.; Lowe, A.; Bickerton, D. Community action across South Australia to secure the future of the Silver Daisy bush in a changing climate. *Australas. Plant Conserv.* **2019**, *27*, 8–10.
36. *EPBC Act 1999. Environmental Protection and Biodiversity Conservation Act*; Commonwealth Government of Australia: Canberra, Australia, 1999.
37. Parliament of South Australia. *National Parks and Wildlife Act*; Parliament of South Australia: Adelaide, SA, Australia, 2014.
38. Department of Environment and Primary Industries. *Advisory list of rare or threatened plants in Victoria—2014*; Department of Environment and Primary Industries: East Melbourne, VIC, Australia, 2014.
39. Peterson, B.K.; Weber, J.N.; Kay, E.H.; Fisher, H.S.; Hoekstra, H.E. Double digest RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE* **2012**, *7*, e37135. [CrossRef]
40. Catchen, J.M.; Amores, A.; Hohenlohe, P.; Cresko, W.; Postlethwait, J.H. Stacks: Building and genotyping loci de novo from short-read sequences. *G3 Genes Genom. Genet.* **2011**, *1*, 171–182. [CrossRef] [PubMed]
41. Catchen, J.M.; Hohenlohe, P.A.; Bassham, S.; Amores, A.; Cresko, W.A. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **2013**, *22*, 3124–3140. [CrossRef]
42. Danecek, P.; Auton, A.; Abecasis, G.; Albers, C.A.; Banks, E.; DePristo, M.A.; Handsaker, R.E.; Lunter, G.; Marth, G.T.; Sherry, S.T. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef]
43. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.; Daly, M.J. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [CrossRef]
44. Meirmans, P.G.; van Tienderen, P.H. GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Mol. Ecol. Notes* **2004**, *4*, 792–794. [CrossRef]
45. Meirmans, P.G. GENODIVE version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Mol. Ecol. Res.* **2020**, *20*, 1126–1131. [CrossRef] [PubMed]
46. Shafer, A.; Peart, C.R.; Tusso, S.; Maayan, I.; Brelsford, A.; Wheat, C.W.; Wolf, J.B. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods Ecol. Evol.* **2016**, *8*, 907–917. [CrossRef]
47. Jombart, T.; Devillard, S.; Balloux, F. Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet.* **2010**, *11*, 94. [CrossRef] [PubMed]
48. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [CrossRef] [PubMed]
49. Rosenberg, N.A. DISTRUCT: A program for the graphical display of population structure. *Mol. Ecol. Notes* **2004**, *4*, 137–138. [CrossRef]
50. Kopelman, N.M.; Mayzel, J.; Jakobsson, M.; Rosenberg, N.A.; Mayrose, I. Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Res.* **2015**, *15*, 1179–1191. [CrossRef]
51. Pembleton, L.W.; Cogan, N.O.; Forster, J.W. St AMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol. Ecol. Res.* **2013**, *13*, 946–952. [CrossRef] [PubMed]
52. Huson, D.H. SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics* **1998**, *14*, 68–73. [CrossRef]
53. Dutheil, J.Y. *Statistical Population Genomics*; Springer Nature: Basingstoke, UK, 2020.
54. Chang, C.C.; Chow, C.C.; Tellier, L.C.; Vattikuti, S.; Purcell, S.M.; Lee, J.J. Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **2015**, *4*, 7. [CrossRef]
55. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing Website: Vienna, Austria, 2020.
56. Lowe, A.J.; Abbott, R.J. Origins of the new allopolyploid species Senecio cambrensis (asteraceae) and its relationship to the Canary Islands endemic Senecio teneriffae. *Am. J. Bot.* **1996**, *83*, 1365–1372. [CrossRef]
57. Rasband, W.S.; ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA. 1997−2018. Available online: https://imagej.nih.gov/ij/ (accessed on 11 June 2021).
58. Grundy, M.; Rossel, R.V.; Searle, R.; Wilson, P.; Chen, C.; Gregory, L. Soil and landscape grid of Australia. *Soil Res.* **2015**, *53*, 835–844. [CrossRef]
59. Harwood, T.; Donohue, R.; Harman, I.; McVicar, T.; Ota, N.; Perry, J.; Williams, K.; CSIRO. *9s Climatology for Continental Australia 1976–2005: Summary Variables with Elevation and Radiative Adjustment. v1*; Data Collection; CSIRO: Canberra, Australia, 2016.
60. Gallant, J.; Austin, J.; Williams, K.; Harwood, T.; King, D.; Nolan, M.; Mokany, K.; CSIRO. *9s Soil and Landform for Continental Australia Analysis of Biodiversity Pattern: Aggregated from 3s Data. v1*; Data Collection; CSIRO: Canberra, Australia, 2018.
61. Robert, J.H. raster: Geographic Data Analysis and Modeling. R package version 3.4-10. Available online: https://CRAN.R-project.org/package=raster (accessed on 11 June 2021).
62. Oksanen, A.; Blanchet, F.; Kindt, R.; Legendre, P.; Minchin, P.; Hara, R.; Simpson, G. vegan: Community Ecology Package. R package version 2.3-2. Available online: https://CRAN.R-project.org/package=vegan (accessed on 11 June 2021).

63. Hallgren, W.; Beaumont, L.; Bowness, A.; Chambers, L.; Graham, E.; Holewa, H.; Laffan, S.; Mackey, B.; Nix, H.; Price, J. The biodiversity and climate change virtual laboratory: Where ecology meets big data. *Environ. Model Softw.* **2016**, *76*, 182–186. [CrossRef]
64. Jeremy James, V. *All Future Climate Layers for Australia—5km Resolution*; James Cook University: Townsville, Australia, 2012.
65. Thuiller, W.; Lafourcade, B.; Engler, R.; Araújo, M.B. BIOMOD—A platform for ensemble forecasting of species distributions. *Ecography* **2009**, *32*, 369–373. [CrossRef]
66. Elith, J.; Phillips, S.J.; Hastie, T.; Dudík, M.; Chee, Y.E.; Yates, C.J. A statistical explanation of MaxEnt for ecologists. *Diver. Distrib.* **2011**, *17*, 43–57. [CrossRef]
67. Miller, A.D.; Nitschke, C.; Weeks, A.R.; Weatherly, W.L.; Heyes, S.D.; Sinclair, S.J.; Holland, O.J.; Stevenson, A.; Broadhurst, L.; Hoebee, S.E. Genetic data and climate niche suitability models highlight the vulnerability of a functionally important plant species from south-eastern Australia. *Evol. App.* **2020**, *13*, 2014–2029. [CrossRef] [PubMed]
68. Blyth, C.; Christmas, M.J.; Bickerton, D.C.; Faast, R.; Packer, J.G.; Lowe, A.J.; Breed, M.F. Increased Genetic Diversity via Gene Flow Provides Hope for Acacia whibleyana, an Endangered Wattle Facing Extinction. *Diversity* **2020**, *12*, 299. [CrossRef]
69. Rodger, Y.S.; Pavlova, A.; Sinclair, S.; Pickup, M.; Sunnucks, P. Evolutionary history and genetic connectivity across highly fragmented populations of an endangered daisy. *Heredity* **2021**, *126*, 846–858. [CrossRef] [PubMed]
70. Silvertown, J. The evolutionary maintenance of sexual reproduction: Evidence from the ecological distribution of asexual reproduction in clonal plants. *Int. J. Plant Sci.* **2008**, *169*, 157–168. [CrossRef]
71. Hu, A.Q.; Gale, S.W.; Kumar, P.; Saunders, R.M.; Sun, M.; Fischer, G.A. Preponderance of clonality triggers loss of sex in Bulbophyllum bicolor, an obligately outcrossing epiphytic orchid. *Mol. Ecol.* **2017**, *26*, 3358–3372. [CrossRef]
72. Carstens, B.C.; Satler, J.D. The carnivorous plant described as Sarracenia alata contains two cryptic species. *Biol. J. Linn. Soc.* **2013**, *109*, 737–746. [CrossRef]
73. Bell, D.; Long, D.G.; Forrest, A.D.; Hollingsworth, M.L.; Blom, H.H.; Hollingsworth, P.M. DNA barcoding of European Herbertus (Marchantiopsida, Herbertaceae) and the discovery and description of a new species. *Mol. Ecol. Res.* **2012**, *12*, 36–47. [CrossRef]
74. Dauphin, B.; Vieu, J.; Grant, J.R. Molecular phylogenetics supports widespread cryptic species in moonworts (Botrychium ss, Ophioglossaceae). *Am. J. Bot.* **2014**, *101*, 128–140. [CrossRef] [PubMed]
75. Van Rossum, F.; Martin, H.; Le Cadre, S.; Brachi, B.; Christenhusz, M.J.; Touzet, P. Phylogeography of a widely distributed species reveals a cryptic assemblage of distinct genetic lineages needing separate conservation strategies. *Perspect. Plant Ecol. Evol. Syst.* **2018**, *35*, 44–51. [CrossRef]
76. Whittall, J.B.; Hellquist, C.B.; Schneider, E.L.; Hodges, S.A. Cryptic species in an endangered pondweed community (Potamogeton, Potamogetonaceae) revealed by AFLP markers. *Am. J. Bot.* **2004**, *91*, 2022–2029. [CrossRef]
77. Christmas, M.J.; Jones, J.C.; Olsson, A.; Wallerman, O.; Bunikis, I.; Kierczak, M.; Peona, V.; Whitley, K.M.; Larva, T.; Suh, A. Genetic barriers to historical gene flow between cryptic species of alpine bumblebees revealed by comparative population genomics. *Mol. Biol. Evol.* **2021**. [CrossRef] [PubMed]
78. Edwards, C.E.; Tessier, B.C.; Swift, J.F.; Bassüner, B.; Linan, A.G.; Albrecht, M.A.; Yatskievych, G.A. Conservation genetics of the threatened plant species Physaria filiformis (Missouri bladderpod) reveals strong genetic structure and a possible cryptic species. *PLoS ONE* **2021**, *16*, e0247586. [CrossRef] [PubMed]
79. Abeyrama, D.K.; Dempsey, Z.W.; Ryan, P.G.; Burg, T.M. Cryptic speciation and population differentiation in the yellow-nosed albatross species complex. *Conserv. Genet.* **2021**, 1–10. [CrossRef]
80. Bickford, D.; Lohman, D.J.; Sodhi, N.S.; Ng, P.K.; Meier, R.; Winker, K.; Ingram, K.K.; Das, I. Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* **2007**, *22*, 148–155. [CrossRef]
81. Vitorino, L.C.; Reis, M.N.O.; Bessa, L.A.; Souza, U.J.B.D.; Silva, F.G. Landscape and Climate Influence the Patterns of Genetic Diversity and Inbreeding in Cerrado Plant Species. *Diversity* **2020**, *12*, 421. [CrossRef]
82. Chala, D.; Brochmann, C.; Psomas, A.; Ehrich, D.; Gizaw, A.; Masao, C.A.; Bakkestuen, V.; Zimmermann, N.E. Good-bye to tropical alpine plant giants under warmer climates? Loss of range and genetic diversity in Lobelia rhynchopetalum. *Ecol. Evol.* **2016**, *6*, 8931–8941. [CrossRef] [PubMed]
83. Ortego, J.; Riordan, E.C.; Gugger, P.F.; Sork, V.L. Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Mol. Ecol.* **2012**, *21*, 3210–3223. [CrossRef]
84. Riordan, E.C.; Gugger, P.F.; Ortego, J.; Smith, C.; Gaddis, K.; Thompson, P.; Sork, V.L. Association of genetic and phenotypic variability with geography and climate in three southern California oaks. *Am. J. Bot.* **2016**, *103*, 73–85. [CrossRef]
85. Christmas, M.J.; Biffin, E.; Breed, M.F.; Lowe, A.J. Finding needles in a genomic haystack: Targeted capture identifies clear signatures of selection in a nonmodel plant species. *Mol. Ecol.* **2016**, *25*, 4216–4233. [CrossRef]
86. Guerin, G.R.; Wen, H.; Lowe, A.J. Leaf morphology shift linked to climate change. *Biol. Lett.* **2012**, *8*, 882–886. [CrossRef]
87. Baruch, Z.; Christmas, M.J.; Breed, M.F.; Guerin, G.R.; Caddy-Retalic, S.; McDonald, J.; Jardine, D.I.; Leitch, E.; Gellie, N.; Hill, K. Leaf trait associations with environmental variation in the wide-ranging shrub Dodonaea viscosa subsp. angustissima (Sapindaceae). *Austral Ecol.* **2016**, *42*, 553–561. [CrossRef]
88. Hill, K.E.; Guerin, G.R.; Hill, R.S.; Watling, J.R. Temperature influences stomatal density and maximum potential water loss through stomata of Dodonaea viscosa subsp. angustissima along a latitude gradient in southern Australia. *Aust. J. Bot.* **2015**, *62*, 657–665. [CrossRef]

89. Wright, I.J.; Dong, N.; Maire, V.; Prentice, I.C.; Westoby, M.; Díaz, S.; Gallagher, R.V.; Jacobs, B.F.; Kooyman, R.; Law, E.A. Global climatic drivers of leaf size. *Science* **2017**, *357*, 917–921. [CrossRef]
90. Meier, I.C.; Leuschner, C. Leaf size and leaf area index in Fagus sylvatica forests: Competing effects of precipitation, temperature, and nitrogen availability. *Ecosystems* **2008**, *11*, 655–669. [CrossRef]
91. Guerin, G.R.; Andersen, A.N.; Rossetto, M.; van Leeuwen, S.; Byrne, M.; Sparrow, B.; Rodrigo, M.; Lowe, A.J. Consistent sorting but contrasting transition zones in plant communities along bioclimatic gradients. *Acta Oecol.* **2019**, *95*, 74–85. [CrossRef]
92. Ottewell, K.M.; Bickerton, D.C.; Byrne, M.; Lowe, A.J. Bridging the gap: A genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. *Divers. Distrib.* **2016**, *22*, 174–188. [CrossRef]
93. Nevill, P.G.; Tomlinson, S.; Elliott, C.P.; Espeland, E.K.; Dixon, K.W.; Merritt, D.J. Seed production areas for the global restoration challenge. *Ecol. Evol.* **2016**, *6*, 7490–7497. [CrossRef] [PubMed]

Chapter 5. General Discussion

Summary

In 'Genomics and the challenging translation into conservation practice', Shafer et al. (2015) concluded that one of the key roadblocks to implementing genomics into conservation practice is the lack of incentive for academic researchers to engage with practitioners and apply their science for practical outcomes. I started my PhD with an aim to do exactly this – provide a link between science and practice and I feel that this has been the primary success of my PhD. I started with a background in practical conservation and restoration, but with minimal experience in applying genetics to conservation issues. Over four years, I worked alongside practitioners and academic researchers to develop management plans for three native Australian plant species, each with a unique set of management questions and challenges. For each body of work, the science has already been applied into practice (Hurren et al., 2019, Faast et al. 2019, Blyth & Breed 2018a, Blyth & Breed 2018b, Blyth et al., 2017), and for me this is the proudest achievement of my research. I developed an understanding of the tools required to analyse and interpret genomic data, the limitations of these data, and how best to use it to suit the needs and requirements of practitioners who may have limited understanding of genetics.

The strength of the research in this thesis lies in the extensive, genome-wide and range-wide sampling of three plant species with contrasting life histories traits (e.g., lifespan, range, pollination mechanism), which have traditionally been used as some of the key predictors of genetic structure and diversity (Broadhurst et al., 2017, Hamrick and Godt, 1996, Hamrick et al., 1992, Lowe et al., 2018). While some of the management goals differ between these species (e.g., genetic rescue of an endangered species to delineating seed zones for a widespread, abundant species), they all rely on understanding core tenets of conservation and restoration genetics (Breed et al., 2019). For this thesis, I used three case studies to answer questions about genetic structure,

genetic discontinuities, levels of genetic diversity and inbreeding in populations, and identifying signatures of selection. I then applied the answers to developing conservation and restoration management plans. While this body of work successfully addresses many of these key questions, perhaps most importantly, it highlights the utility of genomic tools for identifying species-wide population genetic structure, characterising current and historical gene flow, and addressing how genetic diversity can change in response to demographic and environmental change. I found a recurring theme of 'cryptic' genetic breaks within species, and this suggests that 'rules of thumb' based on life history traits may be unrealistic for the conservation of plant species and perhaps not suitable for practice.

I also had several challenges and setbacks during my PhD. I found it difficult to investigate adaptation and selection with reduced representation data and no reference genome. I also tried, and failed, to link the genotype with the phenotype in Chapter 3, experiencing the failure of two common garden trials over two years (report available in Appendix). While this was disheartening, it also provided another clear justification for the use of genomics as a tool to make decisions quickly which, given the current rapid rate of climate change, is increasingly important.

Thesis overview

In this thesis, I conducted range-wide sampling of three native Australian plant species and generated novel population-level reduced representation genomic datasets for each. I used these three studies to explore the use of genomics as a tool with which to address questions about genetic structure, diversity, and adaptation in order to develop management guidelines based on the principles of conservation and restoration genetics. I worked directly with practitioners to ensure that these results were translated into practice. Each study unearthed its own distinct story and a unique set of management recommendations. Here, I summarise the main findings and

discuss them in the context of the fields of restoration and conservation genomics (see Figure 1. For graphical overview).

Chapter 1. Is a literature review of the fields of conservation and restoration genomics conducted at the time of starting my thesis (2017). Aspects of this piece of work were included in the opinion article in Nature Reviews Genetics — '*The potential of genomics for restoring ecosystems and biodiversity*' (Breed et al., 2019).

Chapter 2. For *Acacia whibleyena,* one of Australia's most endangered wattle species, I found extremely high genetic structure amongst isolated stands and low genetic diversity within stands — a species in an "extinction vortex" (Frankham, 2005). The genomic data revealed that a non-local individual was included in a translocated population. I identified that a high proportion of naturally-occurring seedlings was a mix between the local and non-local genotypes in this population, leading to increased genetic diversity – a potential, albeit accidental, 'genetic rescue'. This allowed development of a management plan with local practitioners which would safely explore assisted inter-stand gene flow as a management strategy.

Chapter 3. For *Maireana sedifolia*, a long-lived and keystone arid-zone shrub with a large, continuous range, I utilised population-level genomic sequencing to develop the best seed sourcing advice for postmine site rehabilitation in South Australia. Vast quantities of seed are required for revegetation projects, yet the species produces seed unpredictably, making it a challenging species for restoration purposes. I found two genetic clusters with extremely low genetic structure and similar levels of genetic diversity and inbreeding within each cluster, suggesting strict local seed sourcing was not essential. These results allowed development of a seed sourcing guideline which gave practitioners the best chance of securing enough seed without risking crossing genetic breaks. I did, however, also discover a break in gene flow that could not be explained by obvious biogeographical barriers in the region, and interesting admixture between the two genetic clusters. Attempts to explain

which environmental factors may be shaping the genetic diversity within the species fell short, highlighting the limitations of NGS data in a species without a reference genome. However, I found significant environmental differences between the two genetic clusters, and between those and the admixed populations, suggesting that environment likely plays a role in shaping the genetic differentiation that I found, thus adding weight to the decision not to move seed between genetic clusters until this has been explored further.

Chapter 4. The study of *Olearia pannosa* and its two described subspecies (*O. pannosa* subsp. *pannosa* and *O. pannosa* subsp. *cardiophylla*) unearthed a complex genetic story which fundamentally highlights the importance of using genetics to establish species and subspecies boundaries alongside traditional morphological. The discovery of a potential endangered cryptic species will hopefully lead to conservation intervention for this unique genetic entity. All other samples fell into two overarching genetic clusters with some substructure and admixture in between, and this aligned with leaf phenotype differences. However, these results did not align well with current subspecies boundaries. As in Chapter 2., I found a genetic break between these two genetic groups that cannot be explained by an obvious biogeographic barrier. Additionally, I looked for genotype-environment associations in *O. pannosa* subsp. *pannosa* populations and found that maximum temperature is likely to be an important agent of selection which needs to be considered in conservation strategies.

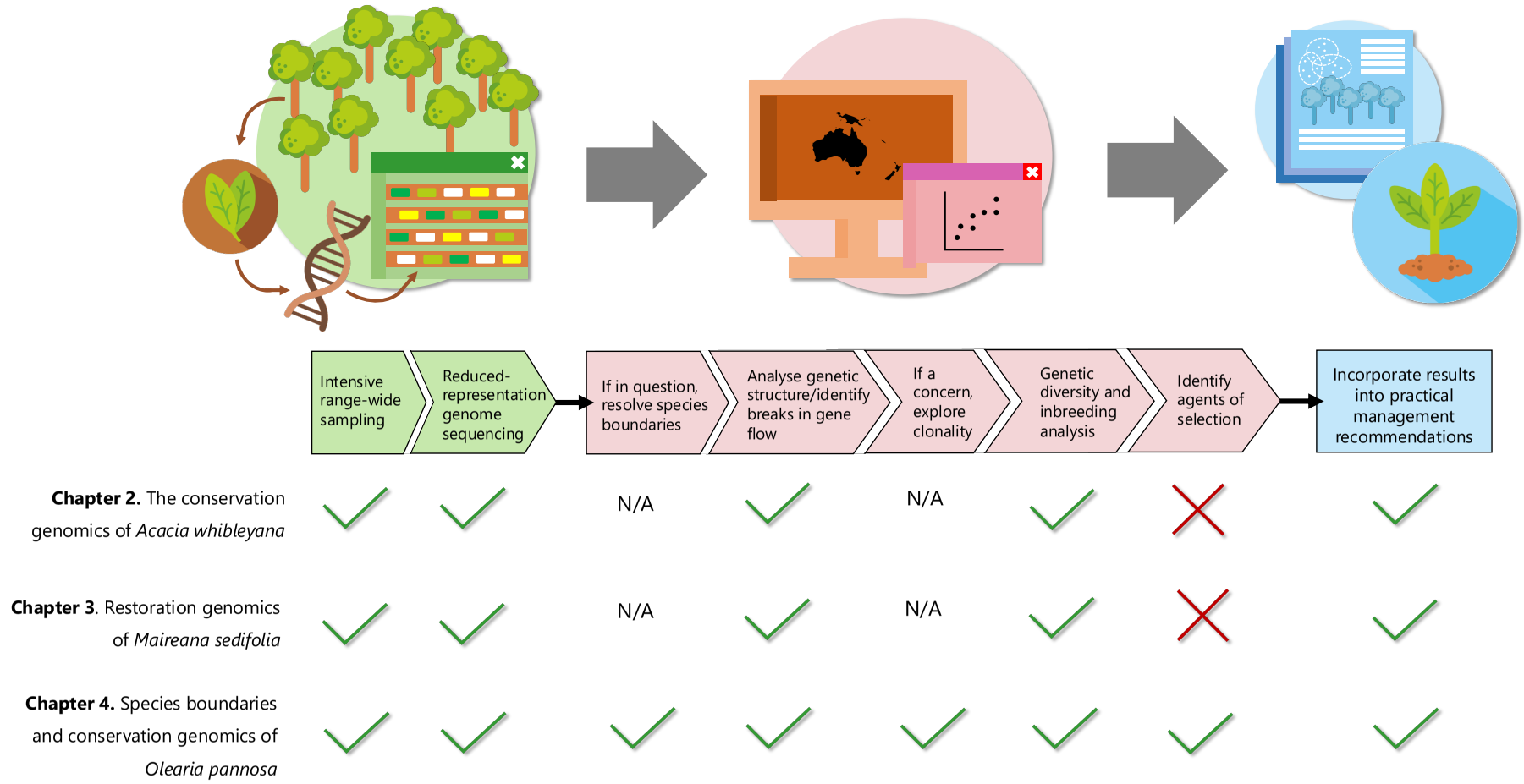


Figure 1. Modified from the workflow for conservation genomic studies of threatened plant taxa by Rossetto et al. (2021). Where appropriate, each component of the workflow was attempted for each species. The relative success of the incorporation of the method is marked with a tick (successful), a cross (unsuccessful), or N/A (not applicable).

In the four years since starting my PhD, I feel that the discourse over using genomic data to guide seed sourcing has shifted from justifying the use of genomic data and heralding the possibilities (Breed et al., 2019, Funk et al., 2012, Luikart et al., 2003, Williams et al., 2014), to focussing on implementing this data to guide how to manage plant species for conservation and restoration (Bell et al., 2019, Breed et al., 2019, Rico, 2021, Rossetto et al., 2019, Rossetto et al., 2021), along with publications which implement these methods, including my own outputs from this doctoral thesis (Blyth et al., 2021, Blyth et al., 2020, Broadhurst et al., 2021, Carvalho et al., 2021, Rossetto et al., 2020). Yet do I feel that the results of my thesis have lived up to the suggestions of the power and promise of genomics? I will now discuss this in the context of each of the central questions that I hoped genomics would answer.

Taxonomy and genetic differentiation

A logical first step when developing conservation and restoration programmes is deciding which species need to be conserved or restored. This can be challenging when questions remain over taxonomy (Supple and Shapiro, 2018). Chapter 4. is a classic example of this scenario and how genomics can be successfully used to address these questions. Funding was acquired to develop a conservation strategy for *Olearia pannosa* but unresolved questions over the taxonomic status of the subspecies needed to be resolved first. Indeed, the importance of using genetic data for such a study is highlighted by the extreme genetic divergence between samples previously thought to belong to the same subspecies, adding to a growing body of evidence that these cryptic species are more commonplace than previously thought (Carstens and Satler, 2013, Christmas et al., 2021, Dauphin et al., 2014, Whittall et al., 2004).

A further use of genetics in seed sourcing studies is identifying gene flow between populations and any genetic breaks that may be present. If there is little gene flow and genetic divergence is high between populations, then caution should be applied to conserve potentially unique genotypes and to avert an outbreeding

depression event through the introduction of new, potentially incompatible alleles. Genetic differentiation can occur due to physical barriers between populations (e.g., a mountain range, habitat fragmentation). Disruption to gene flow can also occur in places with no obvious physical boundaries, and in plant species a common example of this is differing flowering times (e.g., to attract specific pollinators) providing a reproductive barrier. Such barriers to gene flow can be 'cryptic' and are more difficult to predict. For each study species, I found vastly different results in genetic differentiation between populations, with evidence of gene flow across many hundreds of kilometres in Chapter 3, *Maireana sedifolia*, to genetic differentiation in little over a kilometre in Chapter 2, *Acacia whibleyana* — in this case likely due to differences in pollination syndromes between to two species. One of the striking results from this thesis was that for each study, I found evidence of genetic breaks that could not be entirely explained by isolation or obvious biogeographic barriers, and genetically admixed populations at the location of these breaks. This has serious implications for conservation management which I will discuss in the relevant section below.

The benefit of using genome-wide data over methods that focus on only one or a few genetic markers (e.g., microsatellites, AFLPS), is that individuals are more confidently assigned in genetic cluster analyses. For example, Broadhurst et al. (2021) analysed genetic diversity and structure in *Banksia marginata* using both microsatellites and SNP data and compared the results. The genetic structure was far clearer for the SNP dataset, with clear clustering identified using PCA of SNP data but not of the microsatellite data. Furthermore, while the STRUCTURE analyses identified the same likely number of genetic clusters (K), the SNP data assigned individuals to clusters with a far greater degree of confidence (99-100%) while the microsatellite data resulted in many individuals being proportionally assigned to more than one cluster. Having such a strong degree of confidence in cluster assignment is of particular importance for confidently identifying admixed populations and making appropriate

management decisions, and this was one of the benefits of using genomics in conservation identified by Supple and Shapiro (2018). I identified at least one admixed population in each study (between genetic clusters in Chapters 2 and 3, or between subspecies in Chapter 4). The use of SNP data for each thesis chapter has provided a higher level of confidence that these populations are genuinely admixed. This was also explored in Chapter 3. through assessing the confidence intervals of the populations assigned as >20% admixed in the ADMIXTURE analysis and we concluded that we were confident in this assignment. This offers some justification for the uptake of genomics into conservation and restoration.

Genetic diversity and inbreeding

Overall, SNP datasets in all three species were adequate to estimate the levels of genetic diversity and inbreeding. Due to the nature of threatened species, it was not always possible to meet the target of sampling eight individuals per population. I did not manage this for every collection locality in Chapter 4, and this was likely one of the limitations of this study. Recent research has suggested that for genomic datasets 6-8 samples are sufficient to calculate genetic diversity statistics (e.g., H_O, H_E) and that it is possible to adequately calculate genetic differentiation with only two individuals from each population (Nazareno et al., 2017, Willing et al., 2012). Therefore, this could be a notable benefit of using genomic data for rare and threatened species where sampling a set number of individuals is not guaranteed. However, it is worthwhile noting that for each study, I found either admixed populations or, in one instance, a population with a cryptic species. Low sample numbers would have decreased the likelihood of having a good representation of the population. As such, I would still recommend using at least eight individuals from a population if it is possible to do so. I also found that calculating kinship alongside the inbreeding coefficient was a useful metric to include. This is highlighted in the study of *Olearia pannosa* where, for *O. pannosa* subsp. *cardiophylla*, I found low levels of inbreeding but high kinship, and this may have been due to inbreeding avoidance mechanisms in this group. To avoid

missing this information, I would recommend kinship to be a metric included alongside the inbreeding coefficient in plant conservation or restoration studies. This is important because in self-incompatible plant species, seed set has been observed to decrease as kinship increases (Kenrick and Knox, 1989), so this could be a concern if a population identified to have high kinship is also isolated with low gene flow. To address this problem, an intervention such as assisted gene flow through translocations could reduce kinship through increasing outbreeding. It would also be beneficial to monitor post-translocation seed set to establish if this has been a successful intervention method.

Adaptation/defining 'adaptive seed zones'

Firstly, I think that it is important to note here that this component of my PhD was severely impacted by COVID-19. The only member of my supervisory panel trained in these methods was my remote supervisor, Dr Matthew Christmas, who is currently based at Uppsala University in Sweden. The final year of my studies (2020) was supposed to include several months in Europe attending an adaptive genomic workshop followed by a research visit to my supervisor (Dr Matthew Christmas) to work on this component together. As an alternative, I attended a workshop online and Dr Christmas supervised me on this component remotely to the best of his abilities. While useful, these were not a real substitute for in person learning and training, and I feel that the full potential of this component of my thesis was not met.

Nevertheless, the results that I did find suggest that using small SNP datasets (e.g. <15,000 markers) in species which lack a reference genome have limited use for studies of adaptation. The process was attempted for each species, but only achieved in my fourth chapter. For each study, numerous methods were explored to analyse adaptation (e.g., Bayescan outlier analyses, LFMM, PCAdapt, redundancy analyses of genetic and environmental data). I had greatest success with the redundancy analysis of genetic variation within *Olearia pannosa* subsp. *pannosa*, which identified significant

associations between genotype and environment. In the third chapter, correlations between variation in spatial distribution of the populations and environmental variation made it difficult to attribute any of the genetic variation solely to variation in environment. This was despite finding significant differences between the environments of each genetic cluster. The collection stands analysed in my second chapter inhabit such a restricted area that environmental differences among stands are minimal and selection is unlikely to have played much of a role in shaping the current genetic diversity and divergence in this species. While this could have been because divergent selection is not prevalent in these populations – that local land clearance had reduced gene flow and genetic drift was driving the high genetic structure that was found – whole-genome sequencing would likely be required to unlock the story further. So, while small, reduced representation SNP datasets (>15,000 markers) allow us to tackle classic conservation and restoration genetics questions with high resolution, most of the genome remains unsequenced and this leaves uncertainties over information that could be missed. A study exploring the use of RADseq data (used for all studies in this thesis) concluded that it was likely that, for studies of local adaptation, many loci under selection were likely to be missed (Lowry et al., 2017). Furthermore, in species without an annotated reference genome it is difficult to find the functional importance of any SNPs identified as putatively adaptive.

Genetics vs genomics

For each chapter, the SNP data used adequately allowed me to asses levels of genetic structure and differentiation, and genetic diversity and inbreeding, likely at a higher resolution than traditional methods using chloroplast gene sequencing, microsatellites or AFLPs. A direct comparison of results between types of data would be necessary to conclusively claim this, but evidence suggests that genomic data can answer classic questions related to conservation and restoration genetics more accurately and with lower sample numbers than previously used methods (Nazareno

et al., 2017, Willing et al., 2012). However, the limitations of this data became apparent when exploring its usefulness for identifying adaptation and applying this to seed sourcing recommendations.

Conservation decision making

For each species I found 'cryptic' genetic breaks which influenced the management recommendations I made. This helps to cement the justification of using genetic studies to guide delineating seed zones beyond local, at least until there is enough information to make 'rules of thumb'. Nevertheless, this thesis shows that this may not be possible, due to the nature of 'cryptic' boundary(s). As the number of studies published which use genomics to help choose seed provenance continues to increase, it will become clearer whether we will need to use genomics for every case, or whether over time we will be able to confidently model genetic structure and diversity based on things such as landscape, environment, and life history. This approach has been attempted in the past although not necessarily to provide direct seed sourcing guidelines (Broadhurst et al., 2017, Hamrick and Godt, 1996, Hamrick et al., 1992, Lowe et al., 2018). However, as more genetic studies unearth the complex evolutionary history of plant species, creating 'rules of thumb' seems to be increasingly less realistic. For example, a recent exploration of differentiation in genetic structure between five closely related *Acacia* species with an overlap in range found extremely different genetic structure between some of the species which in turn directly affected provenancing strategies (Rossetto et al., 2020). Therefore, I would conclude that from my own experiences, and from results of recent research, this is some time away. Modelling genetic structure and differentiation may need to consider more contributing factors alongside life history traits (e.g., environment, biogeography).

For each species, I found clearly defined genetic structure and variation in the levels of genetic diversity and inbreeding. Theoretically, this should have allowed me

to easily follow '*Bridging the gap: a genetic assessment framework for population-level threatened plant conservation prioritization and decision-making guidelines*' by Ottewell et al. (2016). While I was able use the framework to guide some of the management recommendations, I found that it was, at times, a little simplistic for the results that were generated from a genomic dataset. Therefore, it is recommended that this framework is updated for the genomics era. I would recommend the following considerations.

It was difficult to determine clear 'high' and 'low' thresholds for the inbreeding and genetic diversity statistics. This framework was designed around microsatellite and AFLP markers, and there needs to be an updated consensus on how to interpret these results and what the thresholds should be for genomic data. Firstly, discrepancies between data types can be enormous. For example, in their comparison of microsatellite data and SNP data in the species *Banksia marginata*, Broadhurst et al. (2021) found H_0 ranged from 0.38-0.64 in microsatellites and from 0.12-0.25 in the SNP data across many of the same populations. Understanding the likely differences in results between genetic and genomic studies would be beneficial in the transition of conservation and restoration to genomics. Furthermore, it would be important to understand how genetic diversity and inbreeding estimates vary between different types of commonly used genomic data (e.g., DARTr, RADseq). It would also be highly beneficial to review the results of existing genetic and genomic studies and define 'benchmarks' which are indicative of a genetically healthy population (i.e., adequate levels of genetic diversity and gene flow). Due to differing life history traits, and their known impact on genetic diversity and structure, we may find that these benchmarks may need to be more specific to either these traits (e.g., pollination syndrome) or even genus. Once known, it would then be possible to create workflows accessible to practitioners. Additionally, to facilitate comparisons between species, I would recommend that standardised parameters are used if available. For example, GENODIVE v3.2 (Meirmans, 2020), can be used to calculate *F'*, a standardised

calculation of F_{ST}, which is designed to allow a comparison of genetic differention between studies and/or species.

I recommend that kinship is used as a metric alongside inbreeding. A kinship analysis can be used to estimate how many individuals in a population are possible clones. While clonality may arise as a mechanism to avoid inbreeding, reproducing clonally reduces the capability of populations to evolve to change, and can thus risk maladaptation to future climate scenarios (Tierney et al., 2020). Therefore, if high rates of clonality are detected, methods to reduce this could be explored (e.g., assisted gene flow). Rates of kinship can also assist practitioners in quickly detecting the populations at risk of detrimental effects (e.g., poor seed set) caused by high kinship in species with inbreeding avoidance strategies (Kenrick and Knox, 1989). Furthermore, analysing kinship of seedlings can be indicative of how many matings have been successful – if a high proportion of a seedling generation are siblings, then it could suggest that there are limitations to mating success and that genetic intervention may be desirable.

I also suggest providing advice on the best methods to detect genetic adaptation, and if genetic variation which is identified to be under selection, providing guidance on how to incorporate into management decisions. I would suggest from my experiences during my PhD, that for species without a reference genome, a redundancy analysis of the effect of environmental variables on genetic variation to be the most straightforward method. If climate (e.g., temperature) is identified as an agent of selection, this can be used to justify a climate adjusted seed-sourcing approach, as I did in my fourth chapter for *Olearia pannosa* subsp. *pannosa*. It is also worth recommending that, where possible, a reference genome should be assembled for projects which aim to incorporate genetic adaptation, as it allows the identification of the functional importance of genetic adaptation.

Looking to the future

For each species, this thesis has greatly increased our understanding of the evolutionary history of three native Australian plant species, and results has been used to guide conservation or restoration management (Hurren et al., 2019, Faast et al. 2019, Blyth & Breed 2018a, Blyth & Breed 2018b, Blyth et al., 2017). While designing and establishing seed production areas did not fall under the scope of this research, many of the genetic considerations required for SPA design (e.g., selecting suitable provenances, enhancing genetic diversity through crossing between suitable localities, avoiding mixing likely to cause outbreeding depression) have been answered in these studies and can be directly applied in future SPA design. Furthermore, the techniques used in this thesis can readily be applied to long-term monitoring of the seed produced in SPAs.

While overall I consider my thesis a success, I would recommend several studies which could further enhance our understanding of each species and contribute towards making effective management decisions.

Chapter 2. *Acacia whibleyana*. The results of this chapter suggest that there could be benefits of genetic mixing between populations. Alongside my PhD, I also worked as part of the multidisciplinary team which developed the management plan for *Acacia whibleyena* for the Government of South Australia's Department of Environment and Water. Using the results of my genomic analysis I helped design a pollination trial which tests for both a pollination deficit (which may help explain the low seed production, low recruitment and/or poor vigour of recruits) and for inbreeding/outbreeding depression. Exploring the latter helps inform whether there are negative fitness consequences of maintaining unique genetic subpopulations or – as inferred in Blyth et al., (2020) – that there are likely positive fitness consequences of mixing between the populations. I also was involved in designing the establishment plan of a new, mixed source, translocated population to help test the long-term benefit

of mixing genetic material and monitoring the results. These will hopefully be implemented soon.

Chapter 3. *Maireana sedifolia* – A key remaining knowledge gap is whether the two clusters that I found are under different selection pressures, whether this would negatively impact establishment of seed moved across this break. Full genome sequencing would provide a much higher likelihood of establishing whether adaptation to environment may be playing a part in the genetic break that I found. Furthermore, a reciprocal transplant trial would be an excellent way to explore this and would help to identify any fitness advantages or disadvantages of moving seed across the genetic break.

Chapter 4. *Olearia pannosa* – I would recommend assembling a reference genome and the whole-genome sequencing of individuals from all three of the overarching genetic groups. This will allow us to determine whether there is an adaptive basis to the genetic structure that was found and identify the function of any genes under selection. Alongside Professor Andy Lowe and Dr Matt Christmas, I have secured funding to do this in late 2021 through a partnership with Bioplatforms Australia. Furthermore, Trees for Life (SA) have established a common garden trial containing individuals from every South Australian collection locality, however further funding is required for long-term monitoring.

References

Bell, D.A., Robinson, Z.L., Funk, W.C., Fitzpatrick, S.W., Allendorf, F.W., Tallmon, D.A. and Whiteley, A.R., 2019. The exciting potential and remaining uncertainties of genetic rescue. *Trends in Ecology & Evolution*, *34*(12), pp.1070-1079.

Blyth, C., Christmas, M. J., Bickerton, D. C., Breed, M. F., Foster, N. R., Guerin, G. R., Mason, A. R. & Lowe, A. J. 2021. Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy. *Life,* 11**,** 553.

Blyth, C., Christmas, M. J., Bickerton, D. C., Faast, R., Packer, J. G., Lowe, A. J. & Breed, M. F. 2020. Increased Genetic Diversity via Gene Flow Provides Hope for Acacia whibleyana, an Endangered Wattle Facing Extinction. *Diversity,* 12**,** 299.

Blyth, C & Breed, M. F. 2018. Conservation genomics of the endangered silver daisy bush (Olearia pannosa). Technical Report prepared by The University of Adelaide, on behalf of Trees for Life.

Blyth, C & Breed, M. F. 2018. Informing translocations of the endangered Acacia whibleyana with conservation genomics. Technical Report prepared by The University of Adelaide, on behalf of Department for Environment and Water, and Eyre Peninsula Natural Resources Management Board.

Blyth, C, Christmas, M. C., Gellie, N. G., McDonald, J., Breed, M. F. & Lowe, A. J. 2017. Harnessing population genomics to define rangewide seed zones for Maireana sedifolia, an Australian arid zone specialist. Technical Report prepared by The University of Adelaide, on behalf of Iluka Resources. Ltd.

Breed, M. F., Harrison, P. A., Blyth, C., Byrne, M., Gaget, V., Gellie, N. J., Groom, S. V., Hodgson, R., Mills, J. G. & Prowse, T. A. 2019. The potential of genomics for restoring ecosystems and biodiversity. *Nature Reviews Genetics,* 20**,** 615-628.

Broadhurst, L., Breed, M. F., Lowe, A., Bragg, J., Catullo, R., Coates, D., Encinas-Viso, F., Gellie, N., James, E., Krauss, S. & Byrne, M. 2017. Genetic diversity and structure of the Australian flora. *Diversity and Distributions,* 23**,** 41-52.

Broadhurst, L., Bush, D. & Begley, J. 2021. Managing Genetic Diversity and Representation in Banksia marginata (Proteaceae) Seed Production Areas Used for Conservation and Restoration. *Diversity,* 13**,** 39.

Carstens, B. C. & Satler, J. D. 2013. The carnivorous plant described as Sarracenia alata contains two cryptic species. *Biological Journal of the Linnean Society,* 109**,** 737-746.

Carvalho, C. S., Forester, B. R., Mitre, S. K., Alves, R., Imperatriz-Fonseca, V. L., Ramos, S. J., Resende-Moreira, L. C., Siqueira, J. O., Trevelin, L. C. & Caldeira, C. F. 2021. Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration. *Molecular Ecology Resources,* 21**,** 44-58.

Christmas, M. J., Jones, J. C., Olsson, A., Wallerman, O., Bunikis, I., Kierczak, M., Peona, V., Whitley, K. M., Larva, T. & Suh, A. 2021. Genetic barriers to historical gene flow between cryptic species of alpine bumblebees revealed by comparative population genomics. *Molecular Biology and Evolution*.

Dauphin, B., Vieu, J. & Grant, J. R. 2014. Molecular phylogenetics supports widespread cryptic species in moonworts (Botrychium ss, Ophioglossaceae). *American Journal of Botany,* 101**,** 128-140.

Frankham, R. 2005. Genetics and extinction. *Biological Conservation,* 126**,** 131-140.

Frankham, R., Ballou, J. D. & Briscoe, D. A. 2009. *Introduction to Conservation Genetics,* Cambridge, Cambridge University Press.

Funk, W. C., Mckay, J. K., Hohenlohe, P. A. & Allendorf, F. W. 2012. Harnessing genomics for delineating conservation units. *Trends in ecology & evolution,* 27**,** 489-496.

Hamrick, J. L. & Godt, M. J. W. 1996. Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society B-Biological Sciences,* 351**,** 1291-1298.

Hamrick, J. L., Godt, M. J. W. & Sherman-Broyles, S. L. 1992. Factors influencing levels of genetic diversity in woody plant species. *New Forests,* 6**,** 95-124.

Hurren, A., Breed, M., Blyth, C., Mason, A., Cleeves, R., Lowe, A. & Bickerton, D. 2019. Community action across South Australia to secure the future of the Silver Daisy bush in a changing climate. *Australasian Plant Conservation: Journal of the Australian Network for Plant Conservation,* 27**,** 8-10.

Kenrick, J. & Knox, R. 1989. Quantitative analysis of self-incompatibility in trees of seven species of Acacia. *Journal of Heredity,* 80**,** 240-245.

Lowe, A. J., Breed, M. F., Caron, H., Colpaert, N., Dick, C., Finegan, B., Gardner, M., Gheysen, G., Gribel, R. & Harris, J. B. C. 2018. Standardized genetic diversity-life history correlates for improved genetic resource management of Neotropical trees. *Diversity and Distributions,* 24**,** 730-741.

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F. & Storfer, A. 2017. Breaking RAD: An evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. Wiley Online Library.

Luikart, G., England, P. R., Tallmon, D., Jordan, S. & Taberlet, P. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature reviews genetics,* 4**,** 981-994.

Kenrick, J. and Knox, R.B., 1989. Quantitative analysis of self-incompatibility in trees of seven species of Acacia. *Journal of Heredity*, 80(3), pp.240-245.

Meirmans, P. G. 2020. genodive version 3.0: Easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular ecology resources,* 20**,** 1126-1131.

Mijnsbrugge, K. V., Bischoff, A. & Smith, B. 2010. A question of origin: Where and how to collect seed for ecological restoration. *Basic and Applied Ecology,* 11**,** 300-311.

Nazareno, A. G., Bemmels, J. B., Dick, C. W. & Lohmann, L. G. 2017. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Molecular Ecology Resources,* 17**,** 1136-1147.

Ottewell, K. M., Bickerton, D. C., Byrne, M. & Lowe, A. J. 2016. Bridging the gap: A genetic assessment framework for population-level threatened plant conservation prioritization and decision-making. *Diversity and Distributions,* 22**,** 174-188.

Packer, J.G., Blyth, C., Breed, M., and Faast, R. 2019. Acacia whibleyana: Part 2 – Management Plan, Technical Report prepared by The University of Adelaide, on behalf of Department for Environment and Water, and Eyre Peninsula Natural Resources Management Board.

Prober, S. M., Byrne, M., Mclean, E. H., Steane, D. A., Potts, B. M., Vaillancourt, R. E. & Stock, W. D. 2015. Climate-adjusted provenancing: a strategy for climate-resilient ecological restoration. *Frontiers in Ecology and Evolution,* 3, p65.

Reed, D. H. & Frankham, R. 2003. Correlation between fitness and genetic diversity. *Conservation biology,* 17**,** 230-237.

Rico, Y. 2021. A comprehensive landscape genomics approach for seed sourcing strategies in landscapes under varying degrees of habitat disturbance. Wiley Online Library.

Rossetto, M., Bragg, J., Kilian, A., Mcpherson, H., Van Der Merwe, M. & Wilson, P. D. 2019. Restore and renew: A genomics-era framework for species provenance delimitation. *Restoration Ecology,* 27**,** 538-548.

Rossetto, M., Wilson, P. D., Bragg, J., Cohen, J., Fahey, M., Yap, J.-Y. S. & Van Der Merwe, M. 2020. Perceptions of Similarity Can Mislead Provenancing Strategies—An Example from Five Co-Distributed Acacia Species. *Diversity,* 12**,** 306.

Rossetto, M., Yap, J.-Y. S., Lemmon, J., Bain, D., Bragg, J., Hogbin, P., Gallagher, R., Rutherford, S., Summerell, B. & Wilson, T. C. 2021. A conservation genomics workflow to guide practical management actions. *Global Ecology and Conservation,* 26**,** e01492.

Shafer, A. B., Wolf, J. B., Alves, P. C., Bergström, L., Bruford, M. W., Brännström, I., Colling, G., Dalen, L., De Meester, L. & Ekblom, R. 2015. Genomics and the challenging translation into conservation practice. *Trends in ecology & evolution,* 30**,** 78-87.

Supple, M. A. & Shapiro, B. 2018. Conservation of biodiversity in the genomics era. *Genome biology,* 19**,** 1-12.

Tierney, D.A., Ahrens, C., Rymer, P. and Auld, T.D., 2020. The interaction of clonality, breeding system and genomics for a highly threatened plant species and the management implications. *Biodiversity and Conservation*, *29*(9), pp.3009-3029.

Williams, A. V., Nevill, P. G. & Krauss, S. L. 2014. Next generation restoration genetics: applications and opportunities. *Trends in Plant Science,* 19**,** 529-537.

Willing, E.-M., Dreyer, C. & Van Oosterhout, C. 2012. Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers.

Whittall, J. B., Hellquist, C. B., Schneider, E. L. & Hodges, S. A. 2004. Cryptic species in an endangered pondweed community (Potamogeton, Potamogetonaceae) revealed by AFLP markers. *American Journal of Botany,* 91**,** 2022-2029.

Appendix

List of journal publications during candidature

2021

Minter, M., Nielsen, E. S., **Blyth, C.,** Bertola L. D., Kantar, M., Morales, H. E., Orland, C., Segelbacher G., and Leigh, D. M. What is genetic diversity and why is it important for species survival? Frontiers for Young Minds (2021). Accepted but yet to be published.

Blyth, C., Christmas, M.J., Bickerton, D.C., Breed, M.F., Foster, N.R., Guerin, G.R., Mason, A.R. and Lowe, A.J., 2021. Genomic, Habitat, and Leaf Shape Analyses Reveal a Possible Cryptic Species and Vulnerability to Climate Change in a Threatened Daisy. *Life*, *11*(6), p.553. **(Thesis Chapter 4)**

2020

Blyth, C., Christmas, M.J., Bickerton, D.C., Faast, R., Packer, J.G., Lowe, A.J. and Breed, M.F., 2020. Increased genetic diversity via gene flow provides hope for *Acacia whibleyana*, an endangered wattle facing extinction. *Diversity*, *12*(8), p.299. **(Thesis Chapter 2)**

2019

Breed, M.F., Harrison, P.A., **Blyth, C**., Byrne, M., Gaget, V., Gellie, N.J., Groom, S.V., Hodgson, R., Mills, J.G., Prowse, T.A. and Steane, D.A., 2019. The potential of genomics for restoring ecosystems and biodiversity. *Nature Reviews Genetics*, *20*(10), pp.615-628. **(Thesis Chapter 1 contributed towards this article.)**

Hurren, A., Breed, M., **Blyth, C.**, Mason, A., Cleeves, R., Lowe, A. and Bickerton, D., 2019. Community action across South Australia to secure the future of the Silver Daisy bush in a changing climate. *Australasian Plant Conservation: Journal of the Australian Network for Plant Conservation*, *27*(4), pp.8-10.

2018

Baruch, Z., Jones, A.R., Hill, K.E., McInerney, F.A., **Blyth, C.**, Caddy-Retalic, S., Christmas, M.J., Gellie, N.J., Lowe, A.J., Martin-Fores, I. and Nielson, K.E., 2018. Functional acclimation across microgeographic scales in *Dodonaea viscosa*. *AoB Plants*, *10*(3), pp.0-29.

List of reports during candidature

2019

Faast, R., **Blyth,** C., Breed, M., and Packer, J.G. 2019. *Acacia whibleyana*: Part 1 — Literature Synthesis to Inform Conservation of a Threatened Acacia. Report prepared by The University of Adelaide, on behalf of Department for Environment and Water, and Eyre Peninsula Natural Resources Management Board.

Packer, J.G., **Blyth, C**., Breed, M., and Faast, R. 2019. *Acacia whibleyana*: Part 2 – Management Plan, Technical Report prepared by The University of Adelaide, on behalf of Department for Environment and Water, and Eyre Peninsula Natural Resources Management Board.

N. Gellie, J. C., **Blyth, C.,** Baruch, Z., McDonald J. G., Lowe, A, J. Breed, M,F. 2017. *Maireana sedifolia* rehabilitation efficiency and success. Technical Report prepared by The University of Adelaide, on behalf of Iluka Resources Ltd.

2018

Blyth, C & Breed, M. F. 2018. Conservation genomics of the endangered silver daisy bush (*Olearia pannosa*). Technical Report prepared by The University of Adelaide, on behalf of Department for Trees for Life.

Blyth, C & Breed, M. F. 2018. Informing translocations of the endangered *Acacia whibleyana* with conservation genomics. Technical Report prepared by The University of Adelaide, on behalf of Department for Environment and Water, and Eyre Peninsula Natural Resources Management Board.

2017

Blyth, C., Christmas, M. C., Gellie, N. G., McDonald, J., Breed, M. F. & Lowe, A. J. 2017. Harnessing population genomics to define rangewide seed zones for *Maireana sedifolia*, an Australian arid zone specialist. Technical Report prepared by The University of Adelaide, on behalf of Iluka Resources Ltd.

McDonald J, Breed M.F., Gellie N.G., **Blyth C.,** Lowe A. J. 2017. A life cycle review of *Maireana sedifolia* with a focus on seeding triggers. Technical Report prepared by The University of Adelaide, on behalf of Iluka Resources Ltd.

Grant report to the Nature Foundation of South Australia on the failure of the common garden trial of Chapter 3.

Defining restoration seed zones in the genomics era

Colette Blyth, PhD Candidate, The University of Adelaide

Abstract

A major challenge faced by restoration practitioners is how to achieve successful revegetation outcomes in a rapidly changing world. There is an imminent shift in seed sourcing techniques for restoration, away from strict local provenancing and towards seed sourcing based upon the genetic architecture of restoration species. A principal reason for this shift is that climate change is altering conditions faster than some plants can adapt, creating potential for local provenances to become maladapted. Consequently, there is an urgent need to study the genomic basis of adaptation to climate change, as only with this information is it possible to underpin the restoration of more resilient plant populations to climate change with real data.

Introduction

My research explores the use of genomics as a tool to increase our understanding of plant population structure and adaptation to environmental variables. Central to my PhD is a range-wide study of population structure and adaptation to environmental variables in the arid zone specialist plant, *Maireana sedifolia*. To do this, I attempted to incorporate rangewide genomic analysis with a common garden trial.

Objectives of my PhD research:

1. Define seed transfer zones from neutral genomic data
2. Compare the neutral seed transfer zones with those inferred from adaptive genomic data to define adaptive seed transfer zones
3. Explore the congruence between neutral and adaptive seed transfer zones, and adaptive and functional phenotypic variation to cross-validate the genomic methods with established phenotyping approaches – **I received funding support from Nature Foundation SA for this objective**

The goal of my NFSA funded year was to complete the collection of cuttings from across aridity gradients in both the east and west genetic clusters, and establish a common garden trial. I also planned to measure key functional traits in situ of the individuals from which cuttings were obtained, and pending the establishment of the cuttings, begin part 1 of the common garden trait measurements. I had also planned to complete the analysis of the neutral genomic dataset and begin the analysis of the adaptive genomic dataset.

Progress, methods and results for each component

1. Define seed transfer zones from neutral genomic data

The analysis for this component is complete (Figure 1.). Genetic diversity did not vary greatly across the species range, suggesting that inbreeding is not a concern. We found evidence for two distinct genetic clusters which provided the basis of delineating two separate seed transfer zones, divided into the more arid western zone and the more mesic eastern zone. There is evidence of admixture between the two genetic clusters in the more arid areas of the eastern seed transfer zone which will be further explored in the adaptive analysis, with a focus placed on the design of climate-adjusted provenancing based upon environmental adaptation.

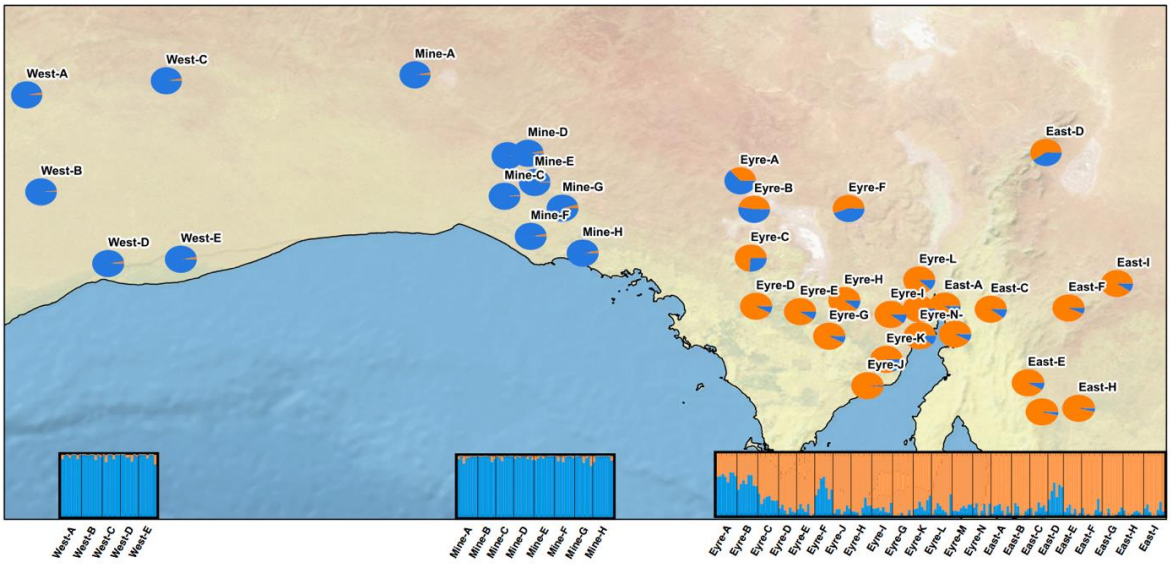


Figure 1: Neutral genetic population structure from a preliminary ADMIXTURE analysis. There have been two clear genetic clusters identified, with an east/west divide. Pie charts at each sampling site show the overall proportion of individuals assigned to the two genetic clusters (blue represents the western cluster, orange represents the eastern cluster).

2. Compare the neutral seed transfer zones with those inferred from adaptive genomic data to define adaptive seed transfer zones

This objective has been impacted by COVID-19 travel restrictions. To complete this objective, I still require further training in the techniques to analyse genomic data for adaptation and selection. I received funding to combine attendance of a highly relevant population genomics workshop in Berlin in May 2020 (https://www.physalia-courses.org/courses-workshops/course9/), with relevant training and work alongside my co-supervisor Dr Matt Christmas (https://katalog.uu.se/profile/?id=N17-740) based at Uppsala University, Sweden to achieve this objective. However, this trip has been cancelled and there has been some delay to the analysis. However, with online resources and collaboration with Dr Matt Christmas, this is still progressing with a projected completion date of 13/7/2021.

As the objective is ongoing, future results will be shared with **Nature Foundation SA.**

3. Explore the congruence between neutral and adaptive seed transfer zones, and adaptive and functional phenotypic variation to cross-validate the genomic methods with established phenotyping approaches – **I received funding support from Nature Foundation SA for this objective**

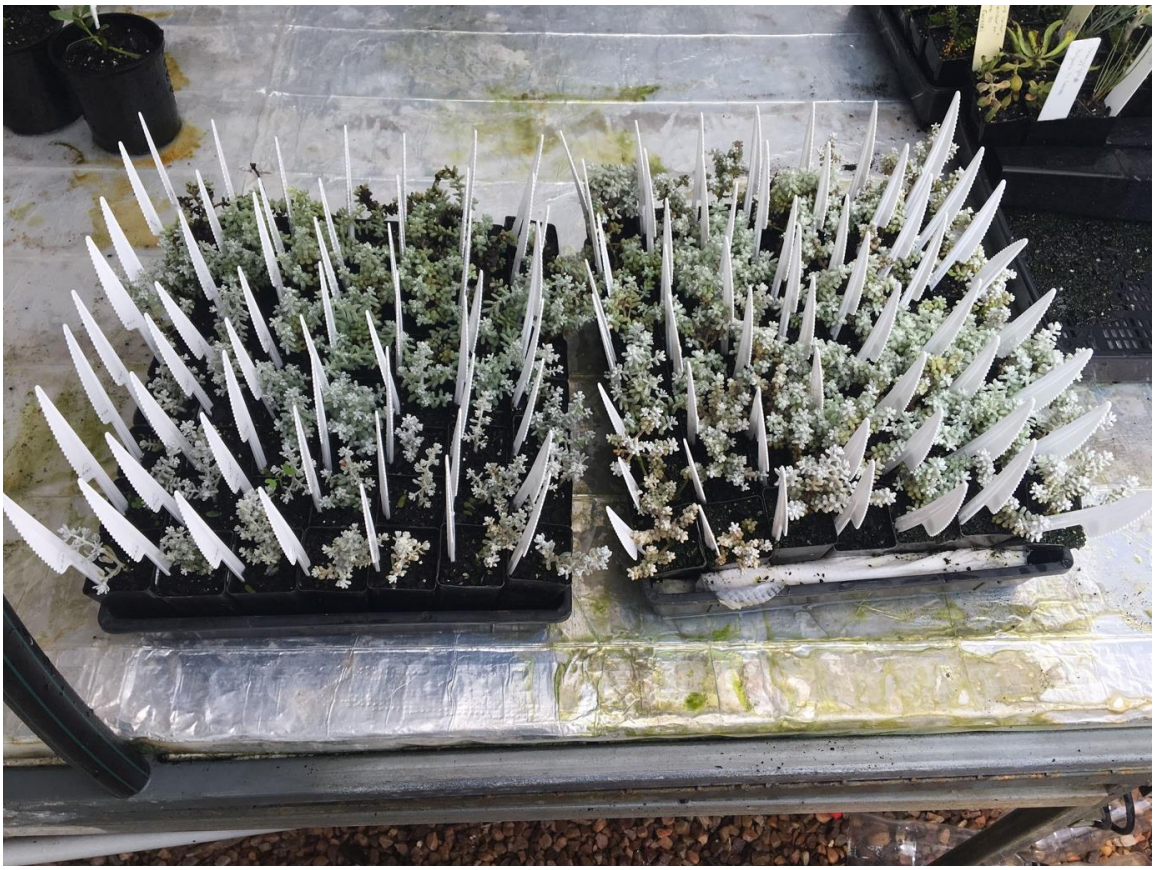
Fieldwork and sample collection
Fieldwork was completed successfully in November 2018. For logistical reasons, the Gawler Ranges were sampled rather than the Nullarbor. Cuttings were taking and propagated from four sites each along two aridity gradients. Further samples were collected from 10 individuals at each site for functional trait measurements

and further molecular analysis (e.g. isotopes) pending the successful establishment of the trial. Fieldwork was completed in less time than originally planned, with the project being completed under budget.

Cuttings and samples being collected

Common Garden Trial

Cuttings were established and kept at the Australian Arid Lands Botanic Gardens, Port Augusta. Throughout my research, they have provided expert advice about *Maireana sedifolia*, as they propagate the species commercially. Unfortunately, establishment of cuttings failed, caused by a heatwave over the 2018/2019 Christmas break, which caused some cuttings to die and some to receive considerable damage. We were informed by horticultural staff at the Australian Arid Lands Botanic Gardens that only 6 cuttings survived, leaving us without the numbers needed for effective sample sizes.



Collected cuttings at the Australian Arid Lands Botanic Gardens

Trait measurements

The initial phase of leaf trait measurements (of the in-situ plants) was completed in December 2018. However, since only 6 cuttings survived the first year of the common garden trial, the decision has sadly been made not to pursue the phenotypic trait component of this research.

Discussion

The failure of the common garden trial caused a great disappointment to the team – a lot of hard work went into the collections and initial trait measurements and the addition of phenotypical data to the genomic dataset would have provided an extra dimension to the project. Fortunately, we can still answer many of the overarching questions of the research with the genomic data. We can still explore adaptation across the species range and answer some of the questions arising over the breaks in gene flow and the admixture that we can see between the genetic clusters. We are currently preparing a manuscript which includes both the neutral and adaptive analyses, with the neutral analyses complete and the adaptive analyses in progress.

We would like to thank **Nature Foundation SA** for their valued contribution to this project. We will notify **Nature Foundation SA** of any updates to this research and will acknowledge them in all resulting publications.