Using ecological and genomic approaches to restore Australian grasslands in the face of global change.

Nissanka Prasangi De Silva

BSc. in Plant Science, University of Colombo

A thesis submitted for the degree of *Doctor of Philosophy* at

Monash University in 2022

School of Biological Science

Copyright notice

Abstract

Australian grasslands are extensively degraded due to land use change and will face further destruction in the future with predicted climate change. Using climate change resilient ecological restoration strategies is important to ensure the long-term protection of these biodiversity rich ecosystems. Introducing non-conventional seed provenencing methods via assisted gene flow could improve restoration outcomes for these ecosystems. For this, knowledge related to the evolutionary potential of candidate native grass species with suitability to be used in grassland restoration is essential. In my thesis I investigate the suitability of two native Australian grass species *Bothriochloa decipiens* and *Bothriochloa macra* for assisted gene flow during restoration by investigating their patterns of population structure and assessing signals of local adaptation to climate using phenotypic and genomic data. For landscape genomic analysis a genome assembly is an important resource. I first present a chromosome-level genome assembly, annotation, and comparative analysis of *B. decipiens*. Comparative analysis revealed that the species is a diploidized allotetraploid. Further I found evidence for biased fractionation and biased retention of duplicated genes following the recent whole genome duplication event. I also discovered an overrepresentation of genes putatively involved in drought response that were retained as duplicated, potentially indicating their involvement in climate adaptation. In a landscape genomic analysis of *B. decipiens* using genotype by sequencing, I discovered evidence for climate adaptation especially for loci related to drought stress response. Further these climate adaptation loci were more likely to be retained as duplicated, supporting a role for allopolyploid in recent climate adaptation. In *B. macra,* using seedling traits measured in common gardens, I found evidence consistent with local adaptation to climate where seedling traits were significantly associated with climate variables. Specifically, I found reduced growth in arid regions which is suggestive of the evolution of a drought tolerance strategy in these *B. macra* populations. However, other seedling traits did not show signatures of divergent selection using an F_{st}-Q_{st} approach. Overall, my thesis includes one of the few genomic studies of an Australian understory species, and findings from my thesis will help in planning climate change resilient restoration management projects in the future.

Publications during enrolment

In Review:

De Silva, N.P., Lee, C., Battlay, P., Fournier-Level, A., Moore, J. L. & Hodgins, K.A. Genome assembly of an Australian native grass species reveals a recent whole genome duplication and biased gene retention of genes involved in stress response.

Thesis including published works declaration.

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes zero original papers published in peer reviewed journals and one submitted publications. The core theme of the thesis is using ecological and genomic approaches to restore Australian grasslands in the face of global change. The ideas, development and writing of all the papers in the thesis were the principal responsibility of myself, the student, working within the School of Biological Sciences, Monash University under the supervision of Dr. Kathryn Hodgins.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.
In the case of chapter 2, 3 and 4 my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s) Nature and % of Co-author's contribution*	Co-author(s), Monash student Y/N*
2	Genome assembly of an Australian native grass species reveals a recent whole genome duplication	Submitted	70% data analysis, writing and editing of paper	1. Lee, C., conducted molecular lab work, analysed data 10% 2. Battlay, P., analysed data 5% 3. Fournier-Level, A., supervised	No

	and biased gene retention of genes involved in stress response			research, manuscript edit 2.5% 4. Moore, J. L., supervised research, manuscript edit 2.5% 5. Hodgins, K.A., conceived study, analysed data, paper writing and editing, supervised research 10%	
3	Signals of climate adaptation and predictions of future maladaptation in an Australian native grass	Not submitted	70% sampling, data analysis, writing and editing of paper	1. Lee, C., conducted molecular lab work 5% 2. Battlay, P., analysed data 10% 3. Fournier-Level, A., supervised research, manuscript edit 2.5% 4. Moore, J. L., supervised research, manuscript edit 2.5% 5. Hodgins, K.A., conceived study, paper writing and editing, supervised research 10%	No

4	Australian native grass *Bothriochloa macra* shows signals of local adaptation to drought	Not submitted	75% sampling, growth chamber experiments, data analysis, writing and editing of paper	1. Lee, C., conducted molecular lab work 5% 3. Fournier-Level, A., supervised research, manuscript edit 5% 4. Moore, J. L., supervised research, manuscript edit 5% 5. Hodgins, K.A., conceived study, paper writing and editing, supervised research 10	No

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student name: Nissanka Prasangi De Silva

Student signature: 

Date: Aug 19, 2022

I hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author, I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor name: Dr. Kathryn Hodgins

Main Supervisor signature: 

Date: Aug 19 2022

Acknowledgements

The successful completion of this thesis would not have been possible without the support of many people. First and foremost, my gratitude goes to my supervisor Dr. Kathryn Hodgins for giving me this amazing opportunity to come to Australia and be a part of her lab. Kay, thank you so much for your patience and understanding especially during the most stressful and rough times I faced during the last four years. Thank you for your guidance throughout this journey and your brilliance in making the best out of all the roadblocks that came in my way. Thank you for your help and direction in bioinformatics and manuscript writing. You encouraged me to look at my best self when I doubted my abilities and for that I am forever grateful.

I would also like to thank my other two supervisors, Dr. Joslin Moore and Dr. Alexandre Fournier-Level for their guidance and help during the journey. Thank you for providing constant feedback on my progress and ideas on how to improve my work.

Thank you to my panel members, Professor Carla Sgro, Professor Paul Sunnucks and Dr. Carly Cook for their advice on the improvement of my thesis. Special thanks to Carla, for her words of wisdom during some hard times.

A big thank you goes to past and present lab members of the Hodgins lab. Thank you, Georgia Loyd, for helping me settle in when I first arrived at Monash and helping me in planning my field collection trips. Thank you, Dr. Hanna Rosinger, for all your advice and tips when I first started, you made me feel welcome. Thank you, Dr. Chris Lee, for helping me with the lab work and teaching me to do molecular lab work. Your patience with me during our lab sessions is much appreciated. Thank you, Dr. Paul Battlay, for all the help in bioinformatics and sharing your code with me.

A special thank you to all the volunteers for allowing me to collect samples in their properties, providing information on places to find plants and also sending me seeds when I could not go and collect them myself.

I would also like to thank my parents and brother for their support in making the transition to a new country and a new life so easy.

Finally, I would like to thank my partner, Angelo, for all his support. You were my field partner through bushfires and global pandemics. You were my therapist when I became overwhelmed during hard times, and you held my hand through thick and thin. Thank you for your patience with me, I know I gave you some really hard times. You are the best. I could not have done it without you.

Table of Contents

List of tables

List of Figures

Chapter 1- General Introduction

Degradation of Australian grasslands

Since European settlement, the clearing of vegetation across Australia has transformed the landscape, fragmenting many ecological communities (Prober et al., 2017). For example, the Box-Gum Grassy Woodland ecological community in south-eastern Australia is federally listed as critically endangered (EPBC Act 1999) Only 10% of this community is left and mainly consists of discontinuous patches of less than 1 hectare (Gibbons & Boak, 2002; Prober & Thiele, 2005). In Southern Australia where much land is used for agriculture, native grassy ecosystems have been heavily impacted as a result of clearing, grazing, pasture improvement and the use of high phosphorus fertiliser and lime (Prober et al., 2002). As a response to this historic tide of destruction the practice of ecological restoration came into practice and in the 1990s the 'decade of land-care' began with large-scale direct seeding of trees and shrubs (Gibbons et al., 2010). A further awakening in implementing restoration efforts will come into practice with the United Nations declaring the start of the decade of restoration in the year 2020.

Ecological restoration in an era of climate change

Ecological restoration can be described as human intervened facilitation of the recovery of a degraded ecosystem primarily due to anthropogenic activities (Rey Benayas et al., 2009). Society for Ecological Restoration Australasia (SERA) defines ecological restoration as the process of assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed. Standards provided by SERA identifies that the term "ecological restoration" is interchangeably used to describe process as well as outcome sought (restored state). These standards favor the use of the term "restoration" for the activity undertaken and the term "recovery" for the sought after or achieved outcome (SERA 2021). SERA also identifies six key principles of ecological restoration practice. Specifically: 1) identification of an appropriate local native reference ecosystem to guide project targets and provide a basis for monitoring and assessing project outcomes; 2) dictating the restoration process after a thorough assessment of the level of resilience and

the degradation of the ecosystem; 3) having restoration targets and goals that are clearly defined with measurable objectives; 4) achieving full recovery of a target site despite the required input investment and time taken; 5) cooperation and knowledge sharing between scientific researchers and restoration practitioners and stakeholders; and 6) carrying out restoration to satisfy not only conservation values but also socioeconomic and cultural values.

Restoration can sometimes be confused with revegetation, but the two processes are recognised as distinct. Regrowth of new vegetation (trees, shrubs, groundcover, and other plants) in disturbed lands usually due to human activities is known as revegetation (Atyeo & Thackway, 2009). Revegetation can be either a natural process occurring due to plant colonisation and succession or can be accelerated artificially by humans. There is a subtle difference between the definitions of revegetation and ecological restoration but revegetating a degraded land with plants that promise a sustainable existence in an unstable future is very important for successful ecological restoration.

Substantial amounts of resources are invested in ecological restoration practices worldwide (Cunningham 2008; Williams et al. 2014). With continuous increase in habitat destruction and human dependence on ecosystems, future costs associated with ecological restoration will intensify globally (Borgström et al. 2016). However, when estimating the cost of restoration, the economic value of the ecosystem services provided by both the intact and the restored ecosystem must be considered (Holl & Howarth, 2000). However, rapid climate change can impact the success of ecological restoration (Harris et al. 2006; Pramova et al. 2019). For example the potential to use ecological restoration as a tool for excess carbon capturing from the atmosphere will be dependent on how climate change influences the regrowth of restored vegetation (Bastin et al. 2019). Therefore the importance of assessing climate change adaptation in ecological restoration practices has been recognized (Heller and Zavaleta 2009; Oliver et al. 2012).

Common practices in ecological restoration and local adaptation

Historically, ecological restoration focused on restoring ecological functions using any

kind of plant, even exotics with many adverse impacts (Prober et al., 2015). However, although now only native germplasm is used to restore sites, many species exhibit extensive functional variation in traits that can impact performance across a landscape. Indeed, the effect of genetic provenance on plant performance has been well documented (Etterson, 2004; Galloway & Fenster, 2000; Jones et al., 2001; Joshi et al., 2001; Linhart & Grant, 1996). Therefore, priority is often given to maintaining local genetic – environment relationships. If a static environment is assumed, local populations are likely best adapted to local conditions, hence leading to the practice of "local provenancing" supported by the idea of "local is the best" (Broadhurst et al., 2008). In other words, the concept of local adaptation, where a local resident population will outperform a foreign genotype in their local environment due to the divergent natural selection occurring in response to the local environment conditions (Kawecki & Ebert, 2004; Savolainen et al., 2013), has been integrated into many management schemes (Hufford & Mazer, 2012; Koch & Samsa, 2007). Meta-analyses have revealed that local populations outperform non-local populations 71% of the time (Hereford, 2009; Leimu & Fischer, 2008), with 45% of the studies indicating the presence of reciprocal fitness benefits (Leimu and Fischer 2008). Consequently, the local-is-best strategy is well supported by empirical evidence. However, if local adaptation is driven by the local environment and the local environment is experiencing rapid change due to climate change, how feasible is it to strictly abide by the "local is the best" paradigm?

Another issue associated with using local seed can be inbreeding depression. It is possible that local populations will be small, subject to strong genetic drift and therefore inbred (Buza et al., 2000; Ellstrand & Elam, 1993; Young et al., 1996). For plants existing as small populations in fragmented landscapes, inbreeding increases either through increased self-pollination or breeding between related individuals (Eckert et al., 2010; Lowe et al., 2005). In small populations inbreeding depression and fixation of deleterious mutations due to reduced efficiency of purifying selection are main contributors to loss of fitness (Lynch et al. 1995; Reed 2003). This decline of fitness through inbreeding could be minimized by genetic purging by removal of deleterious mutations exposed by inbreeding (Leberg and Firmin 2008). Yet studies have shown that purging is rather effective only for moderate sized populations since this purging could come at a demographic cost (García-Dorado 2012). . The issue of inbreeding depression supports the need to practice sourcing seeds from non-local provenances when the size of the local population is diminished.

Alternative seed provenancing strategies for ecological restoration

There has been a more recent push to change from a strict local provenancing strategy towards alternative strategies (Breed et al., 2013; Broadhurst et al., 2008). For example, an effort to mimic natural patterns of gene flow is made with the practice of "composite" provenancing by mixing seeds from local provenances with ones obtained from more distant sites (Broadhurst et al., 2008). "Admixture" provenancing is when seeds are collected from sites across a species range regardless of the whereabouts of the site of interest and mixed (Breed et al., 2013). In "climate adjusted" provenancing a mix of seeds from local genotypes along with sources growing in areas with climates similar to the predicted climate change is included (Prober et al., 2015). All the different three provenancing strategies outlined above are examples of the concept "Assisted Gene Flow".

Concept of "Assisted Gene Flow"

Assisted gene flow is the foundational concept that above non typical provenance strategies are built upon and should be further explored. Assisted gene flow involves in the moving of organisms or germplasm across a current species' range thus transferring allelic diversity to the recipient populations, increasing its potential to adapt to changing conditions (Aitken and Whitlock 2013; Saremi et al. 2019; Weeks et al. 2017). A complementary concept to assisted gene flow is assisted migration. Assisted migration involves the movement of individuals of a species beyond their native species range, when such movement occuring naturally is quite uncommon (Vitt et al. 2010).

Use of assisted migration as a tool for ecological restoration is rather controversial. The ecological risks of introducing species out of their native range are unpredictable and can be harmful (Ricciardi & Simberloff, 2009; Webber et al., 2011). When individuals are removed from extant populations it increases the extinction risks faced by those populations if they are small and vulnerable (Wootton and Pfister 2013). Assisted migration tends to alter the ecosystem related to the recipient population increasing the potential for serious ecological disruptions including extinctions (Bucharova 2017;

Catford et al. 2018). In contrast, the ecological risks of assisted gene flow are relatively low as the species are moved only within their native range and the focal species is already present in the location into which the new individuals are moved (Mueller & Hellmann, 2008).

Managed populations can receive many benefits from assisted gene flow. Assisted gene flow can introduce or increase the frequency of preadapted alleles to new climate conditions, creating a more evolutionary resilient population to climate change regardless of the population size. Inbreeding depression and lack of genetic diversity can cause extinction of small populations (Keller & Waller, 2002). Assisted gene flow can mitigate these effects by adding more genetic diversity and increasing the population size. More local adaptation can be acquired by more genetically varying populations as they have a greater adaptive response to local selection. Alternatively, when the effective population size is reduced, effective local adaptation too will decline due to lack of variation and the greater impact of drift (Blanquart et al., 2012; Yeaman & Whitlock, 2011). Genetic diversity of a population can be increased by intermediate levels of gene flow, in turn facilitating local adaptation (Blanquart et al., 2012; Gomulkiewicz et al., 1999; Kawecki & Ebert, 2004). Apart from increasing adaptive potential, AGF can also improve local fitness (Blanquart et al., 2012; Sexton et al., 2011). However, there are risks associated with introducing genotypes from foreign populations to candidate recipient populations. If ploidy differences and large-scale chromosomal differences between populations are present, assisted gene flow can cause the decline of the resultant population due to outbreeding depression (Edmands, 2007) by causing partial or complete sterility (Frankham et al. 2011). Genetic integrity of the population could be lost and potential genomic extinction can occur during hybridization between distinct populations (Todesco et al. 2016). Alleles carried over by foreign populations can be maladaptive to the local conditions or the combination of them with local alleles may cause outbreeding depression (Edmands 2007). Although AGF requires populations to be genetically differentiated for alleles affecting fitness in different climates, conditions allowing local adaptation to climate may also cause local adaptation to other environmental factors. These other environmental factors might not change over space and time in the same way climate optima does. As a result individuals could be well adapted to new climate

conditions while simultaneously being maladapted to other local environment conditions such as soil type (Wright et al., 2006), local mycorrhizal communities (Kranabetter et al., 2012), photoperiod (Jackson, 2009) or presence of certain pests and pathogens (Van der Putten, 2012). Despite these negative consequences, simulations support the fact that the net benefit for the mean fitness of the population is large when assisted gene flow introduces sufficient alleles for climate adaptation (Aitken & Whitlock, 2013). This suggests that management of populations via AGF should not be prevented unless the populations are facing risks of immediate extinction due to extremely small population size (Aitken & Whitlock, 2013) and other restoration activities should be considered for these populations.

Integration of genomics with ecological restoration

Detecting adaptive genetic differentiation in candidate species is important for practicing AGF. Use of AGF to improve adaptation of a recipient population to a new climate will likely be successful when the recipient population will face future conditions similar to the present conditions of the locally adapted source population. Further adaptation to new climates may require new combinations of genes, allowing selection to act on a diverse genetic base. Common garden studies have been used as the primary means of detecting adaptive differentiation in plants, even resulting in information to parameterize seed transfer guidelines during restoration (Castellanos-Acuña et al. 2018; O'Neill et al. 2014). This approach is used to detect local adaptation among tree populations originating from different climatic regions (De Kort et al., 2013; Savolainen et al., 2007) as well in widespread grasses (Durka et al. 2017; Clair et al. 2013; Beierkuhnlein et al. 2011; López et al. 2020; Galliart et al. 2019; Weston et al. 2021). The rationale behind this method is controlling for phenotypic plasticity and interactions between the genotype and the environment by growing individuals from different populations in a common environment and using quantitative genetic methods to study the bases of complex genetic traits.

Despite the numerous advantages of using common garden experiments, in recent times the study of local adaptation in non-model species has been strongly driven by the study of genetic markers in natural populations (Luikart et al. 2003). This involves sampling tissue from individuals in natural populations, genotyping them using high throughput

methods and conducting genome scan analysis of selection. Screening of genomes for signs of adaptive differentiation is possible with the recent advances in the time and cost-efficient high throughput genome wide DNA sequencing (Allendorf et al., 2010; Morin et al., 2004; Narum & Hess, 2011; Rossetto & Rymer, 2013). Relationships between putative adaptive loci and the environment can be established by using latest sequencing methods combined with climate data and accounting for non-adaptive geographical patterns affecting allele frequencies (Coop et al., 2010; Manel et al., 2010; Orsini et al., 2013).

Integrating genomic, phenotypic and landscape information should be fruitful for finding the signature of adaptive differentiation in phenotypic traits (Hendry, 2013; Jordan et al., 2020; Le Corre & Kremer, 2012; Zhao et al., 2011). Studies have integrated phenotypic and genomic data to detect climate adaptation in many plant species (Camus-Kulandaivelu et al., 2006; Hall et al., 2007; Y. Li et al., 2010). However, false positives (the detection of genomic signatures of climate adaptation not driven by selection) as well as false negatives (the absence of climate associated SNPs when climate adaptation is present) (Lotterhos & Whitlock, 2014) are both possible so critical evaluations of the utility of genomic techniques to assess provenancing strategies under climate change are needed. Indeed, a study of lodgepole found that for decisions relating to assisted gene flow, genotype data are most informative when combined with the phenotypic and climate data (Mahony et al., 2019), but in the absence of phenotypic data, genotypic and climate data can be useful for detecting local adaptation (Mahony et al. 2019; Steane et al., 2014). Use of genome scans to detect local adaptation in other grass species is also reported (Dell'Acqua et al. 2014; Gould et al. 2018). For long lived restoration species setting up new experimental provenance trials may not be time effective with current rate of climate change, therefore population genomic methods could help to provide timely recommendations for climate change resilient ecological restoration (Breed et al., 2019).

Landscape genomics provide insights into evolutionary processes and molecular basis of adaptation helping to understand how species and populations face and adapt to climate change challenges (McKinney et al., 2017). Landscape genomics combines the use of geographic and environmental data with genetic data from a large number of genetic loci to understand how historic climates had shaped the genetic variation of populations in the past (Sork et al., 2013). Genomic analysis that involves non-model organisms has

been hindered by lack of genomic resources such as reference genomes. To do landscape genomic analysis on non-model organisms to generate data to help with restoration decisions, availability of genomic resources such as reference genomes is essential. Reference genomes hold the key to investigate numerous questions essential in species conservation such as demography, inbreeding, hybridization, disease susceptibility, and adaptation (Fuentes-Pardo & Ruzzante, 2017; W. E. Johnson & Koepfli, 2014; Khan et al., 2016; Supple & Shapiro, 2018). Even though reduced representation sequencing is a simple, cost-effective method for generating genome wide marker data, coupling these with data from a reference genome is advantageous in many ways. It improves the reliability of genotype calls (Torkamaneh et al., 2016), will reduce the coverage required for accurate genotyping (Davey et al., 2011), provides for a greater number of markers to improve the inferences gained by subsequent genomic analysis (Shafer et al., 2017) and will allow combining marker annotation with gene information (Gurgul et al., 2019). Another advantage of having access to a reference genome is the ability to characterize genes and gene families that are relevant to species specific conservation (Johnson et al., 2018). A reference genome could provide genomic coordinates for variants facilitating the identification of linked loci (Baetscher et al., 2018). Reference genomes are essential for identification of functional genetic variation underlying phenotypic, fitness and adaptive variation (Capblancq et al., 2020; Hoban et al., 2016; Manel et al., 2016). Therefore these genomic resources when combined with population genomics studies of adaptation have the potential to provide important advances into key evolutionary questions that go beyond those directly related to restoration, including insights related to the genetic basis of adaptation.

Whole genome duplication, genome evolution and adaptation

Polyploidy occurs in two major types: auto and allo - polyploids (Ramsey and Schemske 2003). Autopolyploidy is caused as a result of within species whole genome duplications while allopolyploidy is caused by hybridization events followed by WGD, or by fusion of unreduced gametes from genetically distinct parents (Otto and Whitton 2000). Polyploidy is a common phenomenon throughout the evolutionary history of angiosperms and observed in at least 70% of all species (Grant 2004; Masterson 1994; Otto and Whitton 2000). There have been long discussions on potential adaptive mechanisms related to polyploidy in plant ecology and

evolution (Levin 1983; Ramsey and Ramsey 2014; Soltis et al. 2016) and recently efforts to experimentally evaluate the polyploidy advantage in an ecological context has begun (Maherali et al. 2009; Ramsey 2011; Godfree et al. 2017). Study on the geographic distribution of polyploids (Rice et al. 2019), points out an association between polyploidy prevalence and environmental stress (Van de Peer et al. 2017). There is evidence of how polyploidy helps in improving stress tolerance in plants (Maherali et al. 2009; Zhang et al. 2010; Liu et al. 2011), highlighting the polyploid advantage under a changing environment.

This polyploid advantage could be the result of the increase in the amount of raw genetic material on which evolution can work. Processes such as sub functionalization and neofunctionalization that occur on the genetic redundancy caused through gene duplication can create functional novelty (Ohno 2013; Shiu and Bleecker 2001; Zhang 2003; Blanc and Wolfe 2004; Gout and Lynch 2015). Changes in gene expression and phenotypes that occur as a result of gene duplication through polyploidy is a significant source of evolutionary novelty in plants (Flagel and Wendel 2009).

Study species

Analyses in this thesis were done using two study species *Bothriochloa decipiens* and *Bothriochloa macra*. *Bothriochloa decipiens* (blue pitted grass) is a warm season, perennial, tufted, C4 (a photosynthesis mechanism found in plants growing in dry conditions to minimize water loss, the first carbon product from photosynthesis contains four carbon atoms as opposed to three carbon atoms in C3 plant) grass that can grow up to 1m in height (Stanley & Ross, 1983). The species is known to be frequently cleistogamous (having small inconspicuous closed self-pollinating flowers) in nature (Connor, 1979). Due to its ability to establish well from direct seeding on many soil types, and the ability to withstand pressure caused by overgrazing, it has become an important species for rehabilitation. It is widespread in subtropical NSW and Queensland as well as tropical Queensland (Simon & Alfonso, 2011). It is a close relative, and phenotypically similar to, the polyploid *Bothriochloa macra*. *Bothriochloa decipiens* is part of a cosmopolitan grass genus (Watson, 1992) closely related to *Capillipedium* and *Dichanthium* (together referred to as BCD). A complex history of hybridization and

allopolyploidy is evident in the BCD clade (Estep et al., 2014; Mueller, 2015). *Bothriochloa decipiens* is a diploid member of the group and may be a parental diploid species to some of the present-day allopolyploids in the clade (Sumadijaya 2015). *Bothriochloa macra* is distributed in QLD, NSW, VIC, and SA. It is a perennial tussock grass producing slender reddish flowering stems in summer and early autumn, giving it the common name red grass or red leg grass. Given the fact that this is a C4 species and active in warmer conditions, flowers are produced from summer to early autumn. The species is pollinated by wind and seed dispersal is either by wind or transportation through adhesion. It is also an increaser species, benefiting from grazing and modification of habitat by increasing in abundance making it widespread in overgrazed pastures (Whalley, 1977). Frost tolerance, high drought tolerance, ability to resprout after fire combined with a lifespan of 5-25 years makes it an ideal candidate for ecological restoration of native grasslands. Tolerance to weeds during establishment is also greater in *B. macra* compared to other Australian native grasses (Hagon, 1977). It has the ability to improve good soil properties by accumulation of organic matter (Moore, 1957). All these characters together can make *Bothriochloa. macra* a well-suited potential candidate for the restoration of degraded Australian grasslands.

Along with the two other genera, *Capillipedium* and *Dichanthium*, genus *Bothriochloa* forms a group known as the BCD clade (De Wet & Harlan, 1966). Species in this clade can interbreed even though they are morphologically diverged. This group also has a complex history of hybridization and polyploidy, making this genus important for studying adaptive significance of polyploidy.

Knowledge gap

Grasses are one of the most ecologically important vascular plant groups that make up about 25% of the world's vegetation (Shantz, 1954). However, grasslands worldwide have been under the pressure of degradation due to land use change causing fragmentation of natural populations and erosion of genetic diversity (Harrison et al., 2015). Australian grassland ecosystems are heavily degraded and are poorly conserved (Hobbs & Yates, 2000). Understanding patterns of genetic diversity and evolutionary mechanisms of adaptation to new environments is key for successful conservation of intact grasslands as

well as restoration of degraded systems. Maintaining genetic diversity within a species is often dependent upon many selective and neutral evolutionary processes (Futuyma 2013) and in grasses the evolutionary processes become more complex due to factors such as polyploidy (Cheplick 1998), clonality (van Kleunen et al. 2002) and intrageneric hybridization (Edwards et al. 2010). Due to these complexities, the ability to project findings across different grass species during restoration efforts is limited and this calls for the need of species-specific data and analysis to be present for successful conservation management. However, studies exploring genetic diversity in grasses are mainly aimed at species with agricultural importance (Buckler et al., 2001). Examples of assessing local adaptation for climate related factors in grasses using common garden experiments, measuring phenotypes, and relating the differences in phenotypes to climate is reported in literature (Beierkuhnlein et al., 2011; Faria et al., 2015; Gellesch et al., 2017; Kramer et al., 2018; Kurze et al., 2017; Malyshev et al., 2016; Pirnajmedin et al., 2017; Power et al., 2016). Fewer studies use phenotypic data combined with related genomic data to investigate signals of climate adaptation in ecologically important grass species (Liu et al., 2016; López et al., 2020; Lowry et al., 2013). However a major knowledge gap on the genetic and evolutionary potential of other ecologically important grasses exists. My thesis will not only provide insights into the genetic and evolutionary potential of two ecologically important native Australian grasses but will also generate genomic resources useful for future genomic studies involving native Australian grasses.

Thesis overview

The main aims of my thesis were to investigate the suitability of the two native Australian grass species, *B. decipiens* and *B. macra,* for climate change resilient ecological restoration of grasslands by identifying signatures of local adaptation to climate and to develop genomic resources for these species. In my first data chapter (Chapter 2) I report a chromosome-level genome assembly, annotation, and comparative analysis of *Bothriochloa decipiens*, a species belonging to a group with a complex history of hybridization and polyploidy (BCD clade). This is the first genome assembly and annotation of a species that belongs to this fascinating yet complex group. Findings from this study support hypotheses explaining the biased retention of duplicated genes following polyploidy and point to differences in repeat activity associated with subgenome dominance. The second data chapter (Chapter 3)

presents population structure and landscape genomic analysis of *B. decipiens* populations using genotype by sequencing (GBS). I discovered evidence of significant population differentiation that was structured geographically, as well as rates of self-fertilisation which varied substantially among populations, suggesting local pollen and seed dispersal. Using genotype-environment associations, I discovered evidence of climate adaptation in *Bothriochloa decipiens*. Loci related to drought stress response were overrepresented, suggesting that aridity may be important in structuring adaptive genetic variation in the species. In the third data chapter (Chapter 4) I explore patterns of differentiation of early seed seedling traits of *Bothriochloa macra* in relation to historic climates using phenotypic measures of seedlings of *B. macra* populations across Australia in common gardens. Among the measured traits, total biomass showed significant associations with climate variables related to drought, potentially indicating climate adaptation. Seedlings from high rainfall and low temperature areas yielded higher total biomass values when compared with seedlings from low rainfall and high temperature areas. This is consistent with the hypothesis that populations from more arid regions have evolved a 'drought tolerance' mechanism to survive drought by decreasing resource investment in overall growth and total biomass allocation. In this chapter I also present a population structure analysis *B. macra* using genome wide SNPs generated by GBS data from a smaller set of populations. Evidence of isolation by distance and by the environment was detected, indicating restricted dispersal that may also be influenced by the environment. Chapter 5 includes an overall summary of my findings and discusses my findings in relation to suitability of these two species to be used in future restoration and possible future research directions.

Chapter 2 - Genome assembly of an Australian native grass species reveals a recent whole genome duplication and

biased gene retention of genes involved in stress response.

Nissanka P. De Silva[1*], Christopher Lee[1], Paul Battlay[1], A. Fournier-Level[2], Joslin L. Moore[3] , Kathryn A. Hodgins[1]

[1]School of Biological Sciences, Monash University, Clayton, Victoria, Australia, 3800. [2]School of BioSciences, University of Melbourne, Melbourne, Victoria, Australia, 3010. [3]Arthur Rylah Institute for Environment Research, Heidelberg, Victoria, Australia, 3084. [*]Corresponding author

Manuscript under review in *GigaScience*

Abstract

Background

The adaptive significance of polyploidy has been extensively debated and chromosome level genome assemblies of polyploids can provide insight into this topic. The Australian grass, *Bothriochloa decipiens,* belongs to the BCD clade, a group with a complex history of hybridization and polyploidy. This is the first genome assembly and annotation of a species that belongs to this fascinating yet complex group.

Findings

Using a combination of Illumina short reads, 10X Genomics linked reads and Hi-C sequencing data we assembled a highly contiguous genome of *Bothriochloa decipiens*, with a total length of 1,218.22 Mb and scaffold N50 of 42.637 Mb. Comparative analysis revealed that the species is a diploidized allotetraploid. We clustered the 20 major scaffolds, representing the 20 chromosomes, into the two subgenomes of the parental species using unique repeat signatures. Found evidence of biased fractionation and differences in the activity of transposable elements between the sub genomes prior to hybridization. Duplicates were enriched for genes involved in transcription and response to external stimuli like drought, supporting a biased retention of duplicated genes following whole genome duplication.

Conclusions

Our results support hypotheses explaining the biased retention of duplicated genes following polyploidy and point to differences in repeat activity associated with subgenome dominance. *Bothriochloa decipiens* is a widespread species with the ability to establish across many soil types, making it useful for ecological restoration of Australian grasslands. This reference genome is a valuable resource for future population genomic research involving Australian grasses which may be helpful in ecological restoration projects.

Keywords - genome assembly, annotation, paleo- allopolyploidization, whole

genome duplication (WGD), biased fractionation

Background

Whole genome duplication (WGD), or polyploidy, occurs via the doubling of chromosomal material either involving one species (autopolyploidy) or by hybridization of two species (allopolyploidy). The polyploid origins of many plant species has long been recognized (Grant, 1981; Ledyard Stebbins, 1950), but while polyploidy is commonly observed in angiosperms, its evolutionary importance has been controversial. Some studies support the hypothesis that polyploidy can drive rapid adaptive evolution (Edger et al., 2015; McCarthy et al., 2016), while others have argued that polyploidy has played a minimal role in evolution and contributed little to adaptation (Stebbins, 1971). However, there is growing evidence that ancestral WGD events caused key changes in major angiosperm clades that led to their successful diversification (Soltis & Soltis, 2016).

After a WGD event various molecular changes occur to restore the diploid state, known as diploidization (Doyle et al., 2008; Freeling, 2009; Lim et al., 2008; Mandáková et al., 2010; Tayalé & Parisod, 2013) via genome rearrangement, gene loss and epigenetic change (Doyle et al., 2008; Sémon & Wolfe, 2007). These molecular processes are collectively known as fractionation. Genes that encode DNA repair mechanisms and organelles tend to revert to single copy status following WGD events (Conant, 2014; De Smet et al., 2013). However, other duplicated genes may be retained for a long time and escape deletion or pseudogenization (Freeling et al., 2012). Several theories explain the patterns of duplicated gene retention and their evolutionary fate. For instance, the gene balance hypothesis states that genes coding for products that are dose-sensitive are protected from fractionation because if fractionated, the stoichiometry of the products and other gene products that they interact with will be affected and may bring about negative or lethal effects to the organism (Birchler & Veitia, 2010). These principles are thought to apply to genes responsible for controlling functions related to gene regulation such as transcription factors or kinases acting as hubs with the potential to control entire gene networks (Rody et al., 2017; Tasdighian et al., 2017).

WGD can also give rise to functionally distinct subgenomes. Biased fractionation through the preferential loss of the duplicated genes from the same sub-genome has been observed in many polyploids (Cheng et al., 2012; Renny-Byfield et al., 2017; Thomas et al., 2006; Woodhouse et al., 2010). Gene expression also tends to be biased between retained homeologs, with greater mRNA abundance observed in regions of the genome where gene loss is less common than in corresponding regions (homeologs) with more gene loss (Cheng et al., 2012; Garsmeur et al., 2014; Renny-Byfield et al., 2017; Schnable et al., 2011). The less fractionated and more highly expressed subgenome is referred to as the dominant subgenome, and this asymmetry frequently occurs in hybrids with divergent parental genomes (Edger et al., 2017).

Arguably the most successful plant family in terms of occurrence, ecological dominance and species richness is Poaceae (grasses) (Linder et al., 2018) and approximately 80% of this family are polyploids (Stebbins & Ledyard Stebbins, 1985). Our study species, *Bothriochloa decipiens,* belongs to the tribe Andropogoneae (subfamily Panicoideae). This tribe contains species that have ecological as well as economic importance, and independent allopolyploid events have occurred in exceptionally high numbers in this group (Estep et al., 2014). *Bothriochloa decipiens* is part of a cosmopolitan grass genus (Watson et al., 1992) closely related to *Capillipedium* and *Dichanthium* (together referred to as BCD). These three genera have the ability to interbreed despite their morphological differences, and the term compilospecies was coined to describe this hybrid complex (Harlan & de Wet, 1963; Wet & Harlan, 1970). *Bothriochloa decipiens* is a diploid member of the group and may be a parental diploid species to some of the present-day allopolyploids in the clade (Sumadijaya, 2015).

Here we report a chromosome-level genome assembly, annotation, and comparative analysis of a species in the BCD clade, *Bothriochloa decipiens*. This is the first genome assembly and annotation of a species that belongs to this fascinating yet complex group. Our highly contiguous *B. decipiens* genome assembly showed clear evidence of recent paleo-polyploidy. Using repeat signatures diverged between putative homeologous chromosomes we were able to organise chromosomes into subgenomes, allowing estimation of the timing of the speciation event prior to the most recent allopolyploidy event in this species. We further describe signatures of biased fractionation between

subgenomes, as well as biases in functional annotations of genes retained as duplicated or single copy. This genome reference will act as an important resource for population-genomic analysis of the group and will aid our understanding of the rich history of allopolyploidy in the BCD clade and its evolutionary significance.

Analyses

Genome size estimation, genome assembly and transcriptome assembly

Using flow cytometry (FCM) (see methods) the haploid genome size of the accession COB1- 7 was estimated to be 1.25 Gb (Giga base pairs). The genome assembly was sequenced using a method that combines assemblies from linked read sequencing (10X) with HiRise^{TM} scaffolding achieved through Chicago and Hi-C libraries constructed by Dovetail Genomics (Putnam et al., 2016) (Table 2 - 1). As the final assembly had a L90 represented in 20 scaffolds (Table 2 - 1), we assumed that these scaffolds represented the haploid chromosomes of *B. decipiens* (De Wet & Harlan, 1966; De Wet & Higgins, 1963).

Table 2 - 1. Statistics of the *Bothriochloa decipiens* genome assembly

Assembly statistics	**10X**	**10X+Chicago**	**10X+Chicago+Hi-C**
N50 (size/number)	125 scaffolds / 2.733 Mb	116 scaffolds / 3.080 Mb	10 scaffolds / 53.95 Mb
N90 (size/number)	1,808 scaffolds / 0.023 Mb	616 scaffolds / 0.090 Mb	20 scaffolds / 42.637 Mb
Largest scaffold size	14.847 Mb	14.607 Mb	95.095 Mb

Total number of scaffolds	25,759	19,068	15,895
Total genome size	1,217.63 Mb	1,218.36 Mb	1,218.22 Mb
% Gaps	6.78	6.84	6.86
BUSCO†(*n*)	39:200:13:3	43:200:9:3	45:200:8:2
BUSCO†(%)	15:78:5:1	16:78:3:1	17:78:3:0.8

†Number of BUSCO (Benchmarking Universal Single-Copy Orthologs) genes found in the assembly using the eukaryota odb9 dataset. Genes are split into four categories: complete and single-copy, complete and duplicated, fragmented, and missing, and reported respectively.

The transcriptome is the complete set of transcripts in the cell, and the transcriptome is essential for annotating and interpreting the functional elements of the genome. A transcriptome is essential in the genome annotation process as it helps to determine the transcriptional structure of genes, in terms of their start sites, 5' and 3' ends and splicing patterns. RNA (ribonucleic acid) - sequencing of two tissues (leaf and stem) was used to assemble the transcriptome of *B. decipiens* using Trinity v.2.8.5 (Trinity, RRID:SCR_013048) (Grabherr et al., 2011). The final transcriptome assembly contained 197,655 transcripts, with 104,784 Trinity annotated genes, with an average length of 1,062 bp and a N50 length of 1,677.

Genome annotation, functional annotation, quality validation and repeat identification

We identified 60,652 putative protein coding genes (Table 2- S1) after running iterative runs of MAKER v.3.01.03 (MAKER, RRID:SCR_005309) (Cantarel et al., 2008) genome annotation pipeline that trained gene predictors AUGUSTUS v.3.3.3 (AUGUSTUS, RRID:SCR_008417) (Stanke et al., 2006) and SNAP v.2013-11-29 (Korf, 2004). Functional annotations for the predicted gene annotations were done by searching against several databases (see methods) (Table 2 - 2). We identified 94.1% of the core eukaryotic genes amongst our annotated genes, 22.7% being single copy, 71.4% being duplicated and 2.4% fragmented compared to BUSCO markers present in the library "eukaryota_odb10.2020- 09.10". Of the total assembly length, 52.86% of the genome corresponded to repetitive elements (Table 2 - S2) based on our custom repeat library constructed following recommendations of the MAKERP pipeline for advanced repeat construction (Campbell et al., 2014) (see methods). The majority of the repeated elements were retrotransposons (36.09%) while only 3.34 % were DNA (Deoxyribonucleic Acid) transposons. The two most abundant retrotransposon families were Gypsy and Copia, representing 19.47% and 8.65% of the TEs (transposable elements) identified, respectively (Table 2 - S2). Illustration of the *B. decipiens* genome landscape depicted how gene density was low towards the centre of each scaffold, possibly associated with the centromeres where repeat density was high (Figure 2 - 1).

Table 2 - 2. Summary of functional annotation for protein coding genes.

Database	**Number of gene models with annotations**	**Percentage of gene models with annotations (%)**
UniProtKB/Swiss-Prot	42,417	69.93
Tair10	49,099	80.95
Pfam	43,444	71.62

KEGG (Kyoto Encyclopedia of Genes and Genomes) orthology	17,465	28.79
KEGG orthology - E.C (Enzyme Commission) number annotations	8302	13.68

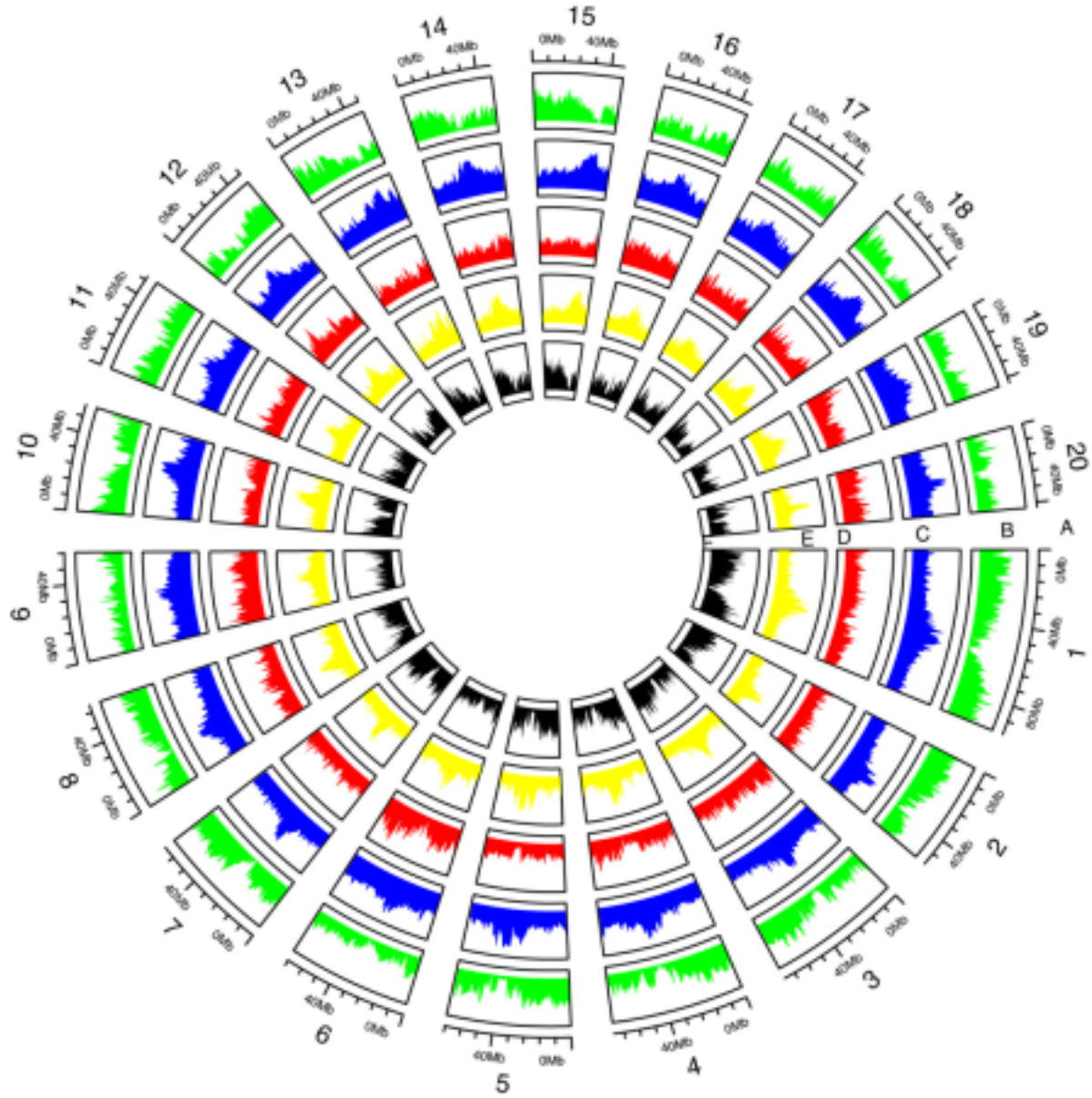


Figure 2 - 1. The *Bothriochloa decipiens* genome landscape. Location across the 20 chromosomes (track A) and distribution, in 1Mb windows of gene density (track B),

repeat density (track C), DNA-TE density (track D), LTR (long terminal repeats) -TE density (track E) and GC (guanine-cytosine) content (track F).

Genome synteny and whole genome duplication

In order to determine if *B. decipiens* had undergone a whole genome duplication, a reciprocal BLASTP (Basic Local Alignment Search Tool Program) (BLASTP, RRID:SCR_001010) (Camacho et al., 2009) was conducted using *B. decipiens* protein sequences as the query against themselves and homeologous scaffolds were identified using the collinear blocks obtained via MCScanX (Wang et al., 2012). This was also corroborated by an alignment of the 20 largest scaffolds (> 40Mb; representing 1.11 Gb of the 1.22 Gb genome) against themselves using Minimap2 v.2.1.8 (Minimap2, RRID:SCR_018550) (Li, 2018) (Figure 2 - S1) . Of these 20 scaffolds, ten pairs (with more than 50% matching across both scaffolds) were identified as the pairs of homeologous scaffolds (Figure 2 - S1). Similarly, we identified collinear blocks between the two putative subgenomes of *B. decipiens* and *Sorghum bicolor*) by conducting a reciprocal BLASTP (BLASTP, RRID:SCR_001010) (Camacho et al., 2009) comparing protein sequences from each species using MCScanX (Wang et al., 2012). Here *Sorghum bicolor* was used for this comparison as it is the closest diploid relative of *B. decipiens* with a high-quality genome assembly and high-quality genome annotations which were essential for the downstream comparative analyses. We identified 33,146 *B. decipiens* genes that were orthologous to 19,611 *S. bicolor* genes across syntenic blocks. A relatively recent paleo polyploidization event was evident as each chromosome from *S. bicolor* almost completely aligned to a pair of *B. decipiens* scaffolds (Figure 2 – 2(B)). Further, these pairs of *B. decipiens* scaffolds show large syntenic blocks of duplicated genes (Figure 2 – 2(A)). There was also evidence of rearrangements between the subgenomes: for example, a translocation from scaffold 18 (homeologous to scaffold 10) to scaffold 8 which appears to have regions from both subgenomes as a result (Figure 2 - 2). Therefore, this translocation likely occurred after the allopolyploidization event. Translocations are apparent in the *B. decipiens* genome alignment against itself (Figure 2 - S1) and also in the syntenic relationship between the *B. decipiens* chromosomes when all 20 of them are aligned against themselves (Figure 2 - S2). Other structural changes can be observed, including several inversions, clearly identifiable on scaffold 13 when

compared to scaffold 15 or to chromosome 6 in *S. bicolor* (Figure 2 - 2 (B)).

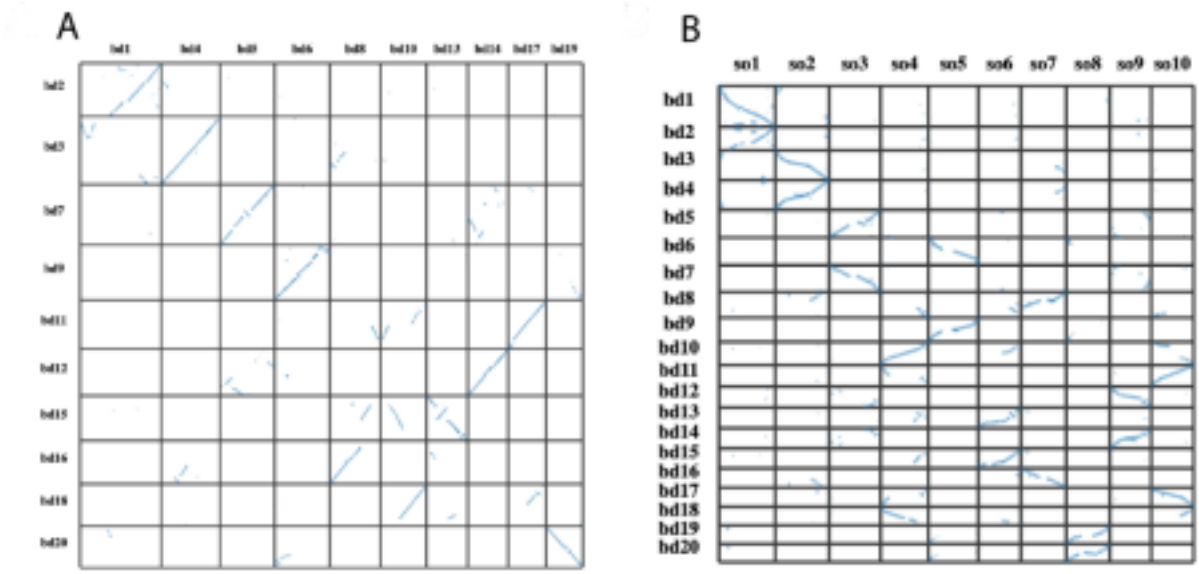


Figure 2 – 2. Plots of genome synteny. (A) The syntenic relationship between the pairs of homeologous chromosomes in *B. decipiens*. (B) The syntenic relationship between the *B. decipiens* and *S. bicolor* orthologous genes. Each *S. bicolor* chromosome (so) shares synteny with a pair of *B. decipiens* chromosomes (bd). This 1:2 relationship suggests that the *B. decipiens* genome was formed by a genome duplication.

Subgenome and homeologous exchange identification

The difference in the distribution of repetitive elements between pairs of homeologous chromosomes can provide evidence of subgenome ancestry (Session et al., 2016). Diagnostic repeat signatures are found on one of each homeologous pair of chromosomes, and they represent the remains of mobile elements with different activity in the diploid ancestors before the merging of the two genomes (Session et al., 2016). As there is no genomic data from any close diploid relatives that do not share the most recent allopolyploidization event, we clustered the putative homeologous chromosomes based on repeat abundance using kmer distributions. We partitioned the *B. decipiens* genome into subgenomes A and B by modifying the methods described in (Mitros et al., 2020) (see methods). We found 919 13- mers (13-bp sequences) occurring at least 100 times across the whole genome and were at least threefold-enriched in one of the homeologous pair relative to the other. Based on the consistent enrichment for these 919 13-mers along the putative homeologous chromosomes, each scaffold of a pair was assigned to a subgenome (Figure 2 - 3). The A group was defined based on the enhanced abundance of 773 13-

mers, and the B group based on the enhanced abundance of the other 146 13-mers (Figure 2 - 3). We then computed the densities of A- and B- preferred 13-mers across the scaffolds (Figure 2 - S3(A)&(B)) and identified potential homeologous exchange between subgenomes. Scaffold 8 from the B subgenome had a high density of subgenome A-preferred kmers at one end of the scaffold (Figure 2 - S3(B)) consistent with the observation of a translocation from the dot plots (Figure 2 - 2). We tested for homeologous exchange among the subgenomes using a Hidden Markov Model implemented (HMM) in the R\HMM package (Himmelmann, 2022). However, we did not find evidence of reciprocal homeologous exchange. Instances of assignment to the alternate subgenome by the HMM occurred in three regions (scaffold 9, 12, and 16), but these were regions of low kmer density, and therefore challenging to assign to subgenomes using this kmer based approach. Also, these instances did not reflect reciprocal exchanges between the subgenomes. Impacts of these ambiguities in subgenome assignment were examined by including and excluding these regions in downstream analyses involving subgenome identification (i.e., biased fractionation). We examined subgenome-enriched LTRs, as differences in LTR activity in parental species can help differentiate subgenomes and can be used to assess the timing of allopolyploidy. We identified 255 LTR repeats that belonged to nine LTR families that were at least three times more common in one subgenome. These repeats also overlapped with A- or B-preferred kmers. There were eight LTR subfamilies (Grande1_ZM_pol/Gypsy, RIRE2_pol/Gypsy, Copia-11_SB/Copia, Copia-73_Mad/Copia, Copia-13_SB/Copia, SZ-7_pol/Gypsy, CRM/Gypsy, Atlantys_OS_polGypsy) identified in subgenome A and only one (Copia-9_SB/Copia) identified in subgenome B. Their genomic locations are shown in Figure 2 - S3(C)&(D).

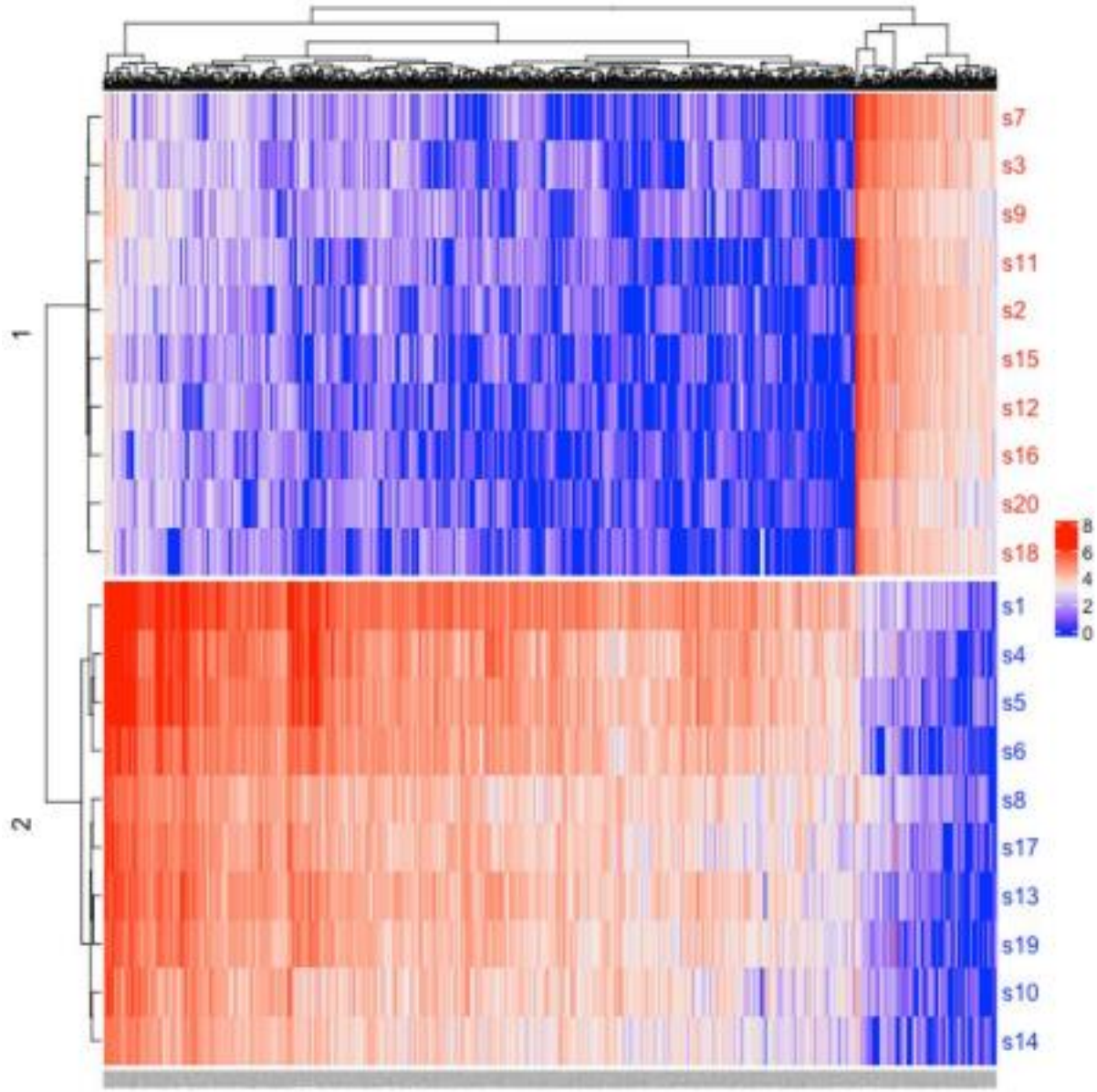


Figure 2 - 3. The differentiation of homeologous pairs of chromosomes into subgenome A (blue) and subgenome B (red) based on the hierarchical clustering of Euclidean distances among scaffolds using counts of 13-mers.

The timeline of paleo tetraploidy

We used 1:1 orthologs identified using OrthoFinder v.2.3.8 (OrthoFinder, RRID:SCR_017118) (Emms & Kelly, 2019) across members of the Andropogoneae tribe (*Sorghum bicolor*, *Zea mays*, *Saccharum spontaneum*, A and B subgenomes of *M. sinensis*). Using *Panicum hallii* and *Setaria italica* as outgroups to identify the

likely timing of the subgenome divergence in *B. decipiens* (see methods) (Figure 2 – 4(A)). We found that the diploid progenitors of the allopolyploid ancestor of *Bothriochloa decipiens* speciated approximately 5.8 MYA (Million years ago). Our tree also dated the divergence of the progenitors of *M. sinensis* to around 6 MYA.

Timing of subgenome-specific LTR expansion

The greater abundance of LTR subfamilies specific to one subgenome is a signal of mobile element activity unique to one of the diploid ancestors of the allopolyploid (Session et al., 2016). Therefore, dating the insertion events of the subgenome-specific LTRs (see methods) can provide a rough estimate of the time that the diploid ancestors of the allotetraploid *B. decipiens* existed independently before the hybridization event. We found that the subgenome-specific LTR activity began 2-3 MYA and peaked around 3.5 MYA (Figure 2 – 4(B)). Activity declined around 6 MYA (Figure 2 – 4(B)), and this coincides with the timing of the allotetraploid speciation inferred through the phylogenetic analysis (Figure 2 – 4(A)). Overall, this suggests that the diploid ancestors may have evolved independently for about 3- 4 MYA before the hybridization event.

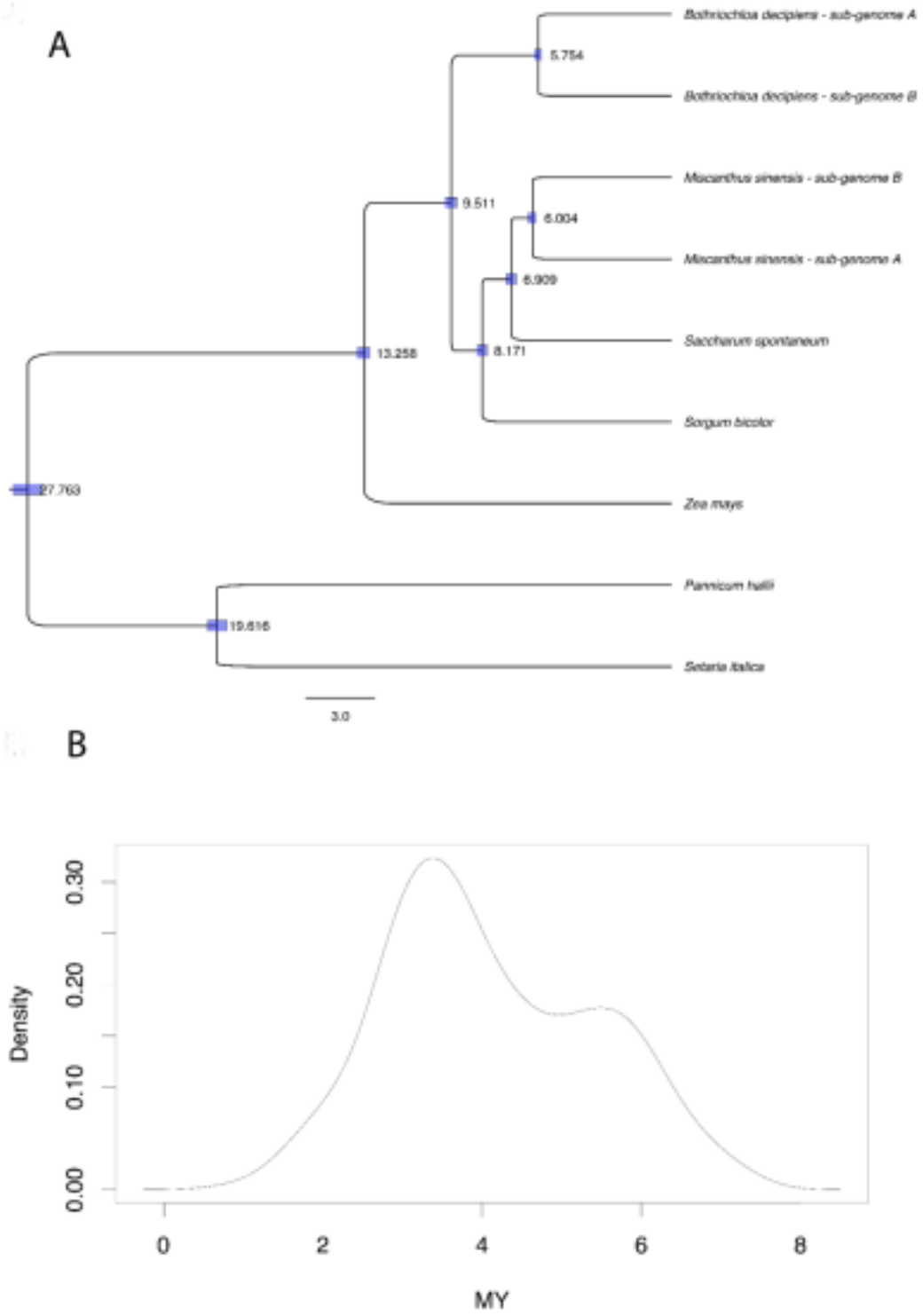


Figure 2 - 4. Phylogeny and sub genome specific LTR expansion. (A) Phylogenetic tree (MYA) of the Andropogoneae showing the time (MYA) of divergence between the diploid ancestors of the allotetraploid *B. decipiens*. (B) Density distribution of divergence the time estimate for all LTR families with sub genome-specific expansion activity.

Biases in gene retention between subgenomes

We analysed the collinear blocks between *S. bicolor* and *B. decipiens* to assess differences in retention of duplicated genes between subgenomes. Subgenome-specific retention was inferred as the number of genes retained in a given subgenome divided by the number of inferred ancestral (i.e., pre duplication) gene numbers. Collinear blocks obtained from the two independent scanning methods McScanX (Wang et al., 2012) and OrthoFinder v.2.3.8 (OrthoFinder, RRID:SCR_017118) (Emms & Kelly, 2019) were used in two independent analyses to confirm any patterns (see methods). We then used a two-sided Fisher's exact test to determine if there was a significant difference in retention of genes between the subgenomes under the null hypothesis that gene loss between sub genomes was random. The percentage of genes retained was higher in subgenome A compared to subgenome B (McScanX, Fisher's exact test, *P*-value = 2.2×10^{-16} and OrthoFinder, Fisher's exact test, *P* value = 1.63×10^{-10} (Table 2 - 3)

Table 2 - 3. Subgenome specific gene retention as observed in analysis with McScanX and OrthoFinder

Clustering Method	**Ancestral genes (retained + single)**	**Ancestral genes retained on A**	**Ancestral genes retained on B**	**Ancestral genes retained on A+B**	**Percent retained on A (%)**	**Percent retained on B (%)**	**Percent retained on A+B (%)**
McScanX	19,523	16,666	15,464	12,195	85	79	62
OrthoFinder	15,262	12,716	12,278	9,620	83	80	53

Differences in the function of duplicated genes retained after the WGD compared to those returning to single copy was tested through gene enrichment analysis using the R\topGo package (topGO, RRID:SCR_014798) (Alexa & Rahnenfuhrer, 2006). All single-copy genes in A or B, or duplicated genes in both A and B were used as foreground genes and the remaining ancestral genes (retained duplicated or single copy) as background genes. GO (Gene ontology) terms related to organelle functions such as chloroplast organisation, chloroplast RNA modification and regulation of mitochondrion organisation were among the top 10 most significant terms associated with genes returning to single copy (Figure 2 – 5(A), Table 2 - S3). Among the ten most significant terms for genes maintained as duplicates were those relating to transcription, including positive and negative regulation of transcription, and terms related to external stress response such as response to water deprivation and salt stress (Figure 2 – 5(B) and Table 2 - S4).

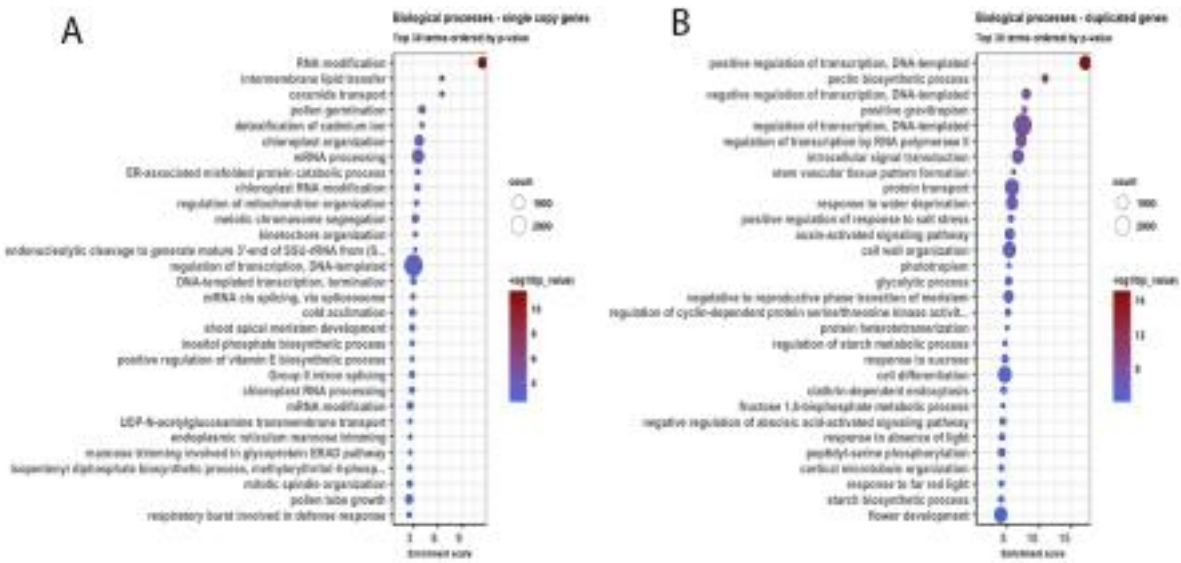


Figure 2 - 5. Summary of the gene enrichment analysis. Top 30 GO terms over-represented in (A) genes retained as single-copy and (B) genes retained as duplicates following the paleo tetraploidy event.

Discussion

Here we report the chromosome-level genome assembly of *Bothriochloa decipiens*, a native Australian grass species important in grassland rehabilitation. Our comparative analysis revealed that this species is a diploidized allotetraploid, consistent with previous phylogenetic analysis of the group (Estep et al., 2014). Our assembly and comparative analysis revealed a relatively recent whole genome duplication. Although the diploid progenitors are unknown, our clustering based on unique repeat signatures grouped the chromosomes into subgenomes. Phylogenetic analysis revealed how the diploid progenitors of the allopolyploid ancestor of *Bothriochloa decipiens* speciated approximately 5.8 MYA. Additionally, we showed evidence of biased fractionation with significantly higher gene retention from one of the subgenomes. This subgenome also appeared to have more active LTRs just prior to the allopolyploidy event. Consistent with hypotheses, genes that were retained as duplicated following the WGD event were enriched for functions involving transcription and stress response.

Patterns of gene loss and retention following allopolyploidy

Biased fractionation commonly occurs after allopolyploidization (Sankoff et al., 2010) and has been observed in *Arabidopsis* (Thomas et al., 2006), maize (Schnable et al., 2009; Woodhouse et al., 2010), and *Brassica* (Cheng et al., 2012), although it is not always observed (Griffiths et al., 2019). Biased fractionation can be a result of genome dominance, where gene expression tends to be higher in one subgenome compared to the other, leading to greater gene loss in the subgenome with reduced expression (Cheng et al., 2012; Garsmeur et al., 2014; Renny-Byfield et al., 2017). Interestingly, subgenome A also appears to have had more active LTRs at the time of the most recent allopolyploidy event, as evidenced by the greater number of diagnostic kmers and LTR families associated with subgenome A (Figure 2 - S3). DNA methylation is known for its role in epigenetic gene silencing (Hirsch & Springer, 2017; Stroud et al., 2013; Weber et al., 2007) and in restricting TE activity (Martienssen & Colot, 2001). It is possible that higher LTR activity prior to the WGD (Figure 2 – 4(B)) and the dominance of subgenome A (Table 2 - 3) are related, and in fact caused by greater genome-wide silencing in

subgenome B. In support of this theory, more retained genes and more active TEs have also been observed in the dominant subgenome of *Miscanthus* (Mitros et al., 2020), but further investigation is warranted.

Functions overrepresented among the genes retained as duplicates, and the genes retained as single copy in the two sub genomes are aligned with theories as well as a large body of empirical work (Blanc & Wolfe, 2004; Duarte et al., 2010; Maere et al., 2005; Paterson et al., 2006; Seoighe & Gehring, 2004). Genes containing domains responsible for functions such as RNA modifications and transmembrane activity tend to revert to singletons in plants (Paterson et al., 2006). Congruently, we identified genes related to RNA modification and transmembrane activity in our GO enrichment analysis of single copy genes (Figure 2 – 5(A)). We also identified enriched GO terms related to organelles such as chloroplast, mitochondria and endoplasmic reticulum in the single-copy gene set. Genes encoding functions related to organelles are commonly retained as single copies (Duarte et al., 2010). An alteration of the dosage balance could explain the retention of single copy genes responsible for organelle-mediated processes such as photosynthesis in chloroplasts and respiration in mitochondria. These functions involve proteins from both the organelle and nuclear genome. The interactions between the two genomes are tightly regulated to maintain the balance of the protein products created from the separate genomes (Haig, 2020; Kleine et al., 2009). During a WGD, this balance could get affected as only the nuclear genome is duplicated, not the organelle's (De Smet et al., 2013). Alternatively, biased fractionation may reduce mixing of genes from the two diploid progenitors (Emery et al., 2018). Nuclear encoded genes performing organelle functions and organelle genes have coevolved in each ancestral genome separately, and biased gene loss and reversion to single copy might maintain coadapted gene complexes and prevent negative interactions of genes between the subgenomes (Emery et al., 2018).

Alongside genes reverting to single copies, many genes were retained as duplicates. Genes coding for subunits of multimeric proteins or complexes, transcription factors and signal transduction mechanisms are biased to avoid fractionation (Blanc & Wolfe, 2004; Maere et al., 2005; Seoighe & Gehring, 2004). GO terms related to transcription, signal transduction and protein biosynthesis were over-represented amongst retained duplicate genes (Figure 2 – 5(B)). The gene dosage hypothesis explains these patterns of biased

retention (Birchler & Veitia, 2010): if either gene copies that codes for a dosage-sensitive gene product that interacts with other gene products is lost, the dosage imbalance may be deleterious to the organism. We also found that retained duplicated genes had an overrepresentation of GO terms related to response to external stressors like water deprivation, salt stress and response to abscisic acid signalling pathways (Figure 2 – 5(B)). Genes that give plants the ability to respond to various environmental stresses are frequently retained as duplicates after WGD (Wu et al., 2020). As neofunctionalization and specialisation can be possible fates of genes retained as duplicates, retention of such genes may promote adaptive evolution to abiotic stresses (Cheng et al., 2018; Defoort et al., 2019; Jiao et al., 2018; Panchy et al., 2016). Multiple paleo polyploidization events that occurred independently throughout the history of evolution of angiosperms could have promoted the diversification of angiosperms across a wide range of environmental conditions by contributing to adaptation to new environments and stress (Zhang et al., 2020).

The evolutionary history of the Andropogoneae

Species in the Andropogoneae clade, which *B. decipiens* belongs to, are dominant in modern day C4 grasslands. Most allopolyploid events in the Andropogoneae occurred recently, in the late Miocene period (Estep et al., 2014), which coincides with the expansion of C4 grasslands (Edwards et al., 2010). Our phylogenetic analysis suggests the speciation event leading to the diploid ancestral genomes occurred at the end of the Miocene, approximately 5.8 MYA, which corresponds to other estimates reported for *Bothriochloa spp*. (Estep et al., 2014). It also appears to have occurred at a similar time as the speciation event leading to the *Miscanthus* A and B subgenomes (Mitros et al., 2020). For *B. decipiens* the insertion times of the subgenome-specific LTRs suggest that these species were diverged for up to 4 million years before the hybridization event, although the date estimates are prone to error due to substitution rate variation among LTR families (Wicker & Keller, 2007).

The genus *Bothriochloa* belongs to a group known as the BCD clade along with the two other genera, *Capillipedium* and *Dichanthium* (De Wet & Harlan, 1966). Species in this clade can interbreed even though they are morphologically diverged. For instance,

Bothriochloa bladhii has been identified as a compliospecies, able to absorb genomes from different species in the BCD complex (De Wet & Harlan, 1966; Estep et al., 2014). Interestingly, a phylogenetic study of species in the BCD clade in Australia suggested that *B. decipiens* may be an ancestral diploid species of the BCD clade, making *B. decipiens* a key species to further our understanding of the evolution of this group (Sumadijaya, 2015). Our high-quality reference genome should spur future comparative genomic studies of allopolyploidy and hybridization in this clade and in grasses more generally.

Most allopolyploidization events in the Andropogoneae are recent, such as the one we report in *B. decipiens*. These events are concurrent with the time of the expansion of C4 grasslands, and some studies argued that the allopolyploidy gave these grasses the ability to adapt to new environments, thereby enabling successful expansion and establishment (Godfree et al., 2017; Linder & Barker, 2014). Future analysis examining the adaptive significance of retained duplicates using both comparative and
population genomic approaches across a greater number of taxa, including groups with frequent WGD such as the Andropogoneae, will further our understanding of the adaptive significance of allopolyploidy and its potential role in niche expansion in C4 grasses.

Potential implications

This genome will be an important resource for population genomic studies involving native grasses in this genus. Such analyses will shed light on the adaptive genetic landscape of these important foundation species, which could play a critical role in the development of climate change resilient grassland restoration practices (Aitken & Whitlock, 2013; Breed et al., 2013). Further this genome will be important in broader comparative analyses of the Andropogoneae which should provide greater insight into the evolutionary significance of allopolyploidy.

Methods

Species description

Bothriochloa decipiens (blue pitted grass) is a warm season, perennial, tufted grass that can grow up to 1m in height (Stanley & Ross, 1983). Due to its ability to establish well from direct seeding on many soil types, and the ability to withstand pressure caused by overgrazing, it has become an important species for rehabilitation. It is widespread in subtropical New South Wales (NSW) and Queensland as well as tropical Queensland (Simon & Alfonso, 2011). It is a close relative, and phenotypically similar to, the polyploid *Bothriochloa macra,* which is a widespread native grass species in south-eastern Australia. The diploid chromosome number of *B. decipiens* is reported to be 2n=40 (De Wet & Higgins, 1963).

Sample collection

The seeds used to grow the diploid *B. decipiens* accession COB1-7 used in this study were collected from Cobbitty, NSW (34°03'N, 150°68'E). Using these seeds, a plant was grown and maintained at Monash University, Clayton to obtain leaf and inflorescence tissue samples for DNA and RNA extractions for the study.

Flow cytometry

We used FCM to estimate the genome size and predict the relative ploidy of 24 populations of *Bothriochloa macra* and *Bothriochloa decipiens,* collected from different locations in states Victoria and NSW. Estimating ploidy from different populations was necessary as *B. decipiens* and polyploid *B. macra* are extremely similar morphologically and ploidy is the best method to reliably distinguish the two species. We also sought evidence for within population variation in ploidy. We estimated the ploidy of at least 5 plants from each population following a modified plant FCM protocol (Dolezel et al., 2007). Leaf samples from each population were collected from greenhouse grown plants and immediately placed on ice for same-day cytometric analysis. Three DNA genome size standards were selected, *Solanum lycopersicum* (2C = 1.96) and *Pisum sativum* (2C = 9.09), and grown from seed. Approximately 40 mg (milligram) of fresh leaf material was

used for each sample, and placed into a 2.0 mL (millilitre) tube with a single 3 mm (millimetre) tungsten carbide bead and 300 µL (microliters) of an ice-cold nuclei suspension buffer modified from de Laats buffer (1984): 15mM (millimolar) HEPES, 1mM EDTA, 0.2% (v/v) Triton X-100, 80mM KCl, 20mM NaCl, 300 mM sucrose, 0.5 mM spermine, 15mM β-mercaptoethanol, 0.25 mM PVP 40. Adjusted to pH7. Samples were placed in a Qiagen Tissuelyser II and ground for 24 seconds at 25 hertz, and then the sample rack was reversed and ground again. The homogenate was filtered through two layers of Millipore Miracloth (22-25 µm (micrometre) pore size) suspended in a 3-piece nozzle. One µL of 10 µg (microgram)/µL RNAse was added for every 100 µL of filtrate and incubated at 37℃ for 20 minutes. Fifteen µL of 0.1 µg/ µL of Propodium Iodide station solution was added to the filtrate and samples were run on the BD Accuri^{TM} C6 Cytometer using the settings outlined in (Galbraith & Lambert, 2012). Internal standards were run on the cytometer at the beginning and end of the session – no change in dye fluorescence was recorded over that period of time. A total of 38 samples produced an observable signal in the FCM run. All samples, excluding standards, were run in a blind fashion so that prior knowledge of expected ploidy did not bias the identification of nuclei peaks. The 2C values were determined for all *Bothriochloa* samples by comparing the FL2-A value of the sample to the internal standards, *Solanum* and *Pisum*, which have a known 2C value of 1.96 and 9.09pg respectively (Dolezel et al., 2007). The average 2C genome size of diploid and polyploid plants was 2.80 pg (range 2.56 –2.99 pg) and 5.38 pg (range 4.94 – 5.91 pg), respectively (Figure 2 - S4). Only one population (COB1) consisted of diploid individuals, and all individuals from this population were tested to confirm our findings. The COB1-7 accession was 2.56 pg which leads to a haploid genome size estimate of 1.25 Gb (Figure 2 - S5). The polyploid samples were likely closely related and phenotypically similar *B. macra*.

DNA extractions

For DNA extraction, fresh leaf tissue was collected from diploid individual COB1-7, flash frozen in liquid nitrogen and stored at -80 ℃. The tissue was then shipped to Dovetail Genomics who completed the DNA extractions. To obtain high molecular weight DNA for 10X Genomics linked read sequencing, 1.8g of leaf material was ground with mortar and pestle to a fine powder to which 200mL of prewarmed CTAB and 100µL BME was

added. This was incubated at 68℃ for 15 minutes. Once incubated, a mixture of 2x phenol chloroform, 1x isoamyl and 0.7x isopropanol was added and centrifuged to form a pellet. The pellet was combined with 9.5 mL of G2, 200µL protease and 19µL RNase. Again, the mixture was incubated at 50℃ for 1 hour. The precipitated genomic DNA was used in library constructions.

10X Library preparation sequencing and 10X assembly

Genomic DNA (gDNA) with an adjusted concentration between 1.0 -1.25 ng/µL was used to prepare the whole genome sequencing libraries using the Chromium Genome Library and Gel Bead Kit v.2, Chromium Genome Chip Kit v.2, Chromium i7 Multiplex Kit and Chromium controller according to manufacturer's instructions (10X Genomics). Genomic DNA was combined with Master Mix, a library of Genome Gel Beads, and partitioning oil to create Gel Bead-in-Emulsions (GEMs) on a Chromium Genome Chip. The GEMs were isothermally amplified with primers containing an Illumina Read 1 sequencing primer, a unique 16bp (base pairs) 10X barcode and a 6bp random primer sequence. Bar-coded DNA fragments were recovered for Illumina library construction. The amount and fragment size of post-GEM DNA was quantified prior using a Bioanalyzer 2100 with an Agilent High sensitivity DNA kit. Prior to Illumina library construction, the GEM amplification product was sheared on an E220 Focused Ultrasonicator (Covaris, Woburn, MA) to approximately 350bp. Then, the sheared GEMs were converted to a sequencing library following the 10X standard operating procedure. The library was quantified by qPCR (quantitative polymerase chain reaction) with a Kapa Library Quant kit (Kapa Biosystems-Roche) and sequenced on a NovaSeq6000 sequencer (Illumina, San Diego, CA) with paired-end 150bp reads.

Chicago library preparation and sequencing

A Chicago library was prepared as described in (Putnam et al., 2016). Briefly, ~500ng of high molecular weight gDNA was reconstituted into chromatin *in vitro* and fixed with formaldehyde. Fixed chromatin was digested with DpnII, the 5' overhangs filled in with biotinylated nucleotides, and then free blunt ends were ligated. After ligation, crosslinks

were reversed and the DNA purified from protein. Purified DNA was treated to remove biotin that was not internal to ligated fragments. The DNA was then sheared to ~350bp fragments and sequencing libraries were generated using NEBNext Ultra II kit with Illumina-compatible indices. Biotin-containing fragments were isolated using streptavidin beads before PCR enrichment of each library. The libraries were sequenced on an Illumina HiSeq X to produce 467 million 2x150bp paired end reads.

Dovetail HiC library preparation and sequencing

A Dovetail HiC library was prepared as described in (Lieberman-Aiden et al., 2009). Briefly, for each library, formaldehyde was used to fix chromatin in the nucleus in place. Fixed chromatin was digested with DpnII, the 5' overhangs filled in with biotinylated nucleotides, and then free blunt ends were ligated. After ligation, crosslinks were reversed and the DNA purified. Purified DNA was treated to remove biotin that was not internal to ligated fragments. The DNA was then sheared to ~350bp mean fragment size and sequencing libraries were generated using NEBNext Ultra enzymes and Illumina-compatible adapters. Biotin-containing fragments were isolated using streptavidin beads before PCR enrichment of each library. The libraries were sequenced on an Illumina HiSeq X to produce 400 million 2x150bp paired end reads.

Genome assembly

The 10X sequence data were assembled *de novo* with Supernova (Supernova assembler, RRID:SCR_016756) (Weisenfeld et al., 2017). This *de novo* assembly, along with Chicago library reads and Dovetail HiC library reads, was used as input data for HiRise, a software pipeline designed specifically for using proximity ligation data to scaffold genome assemblies (Putnam et al., 2016). An iterative analysis was conducted. First, Chicago library sequences were aligned to the draft *de novo* assembly from Supernova using SNAP (Zaharia et al., 2011). The separations of Chicago read pairs mapped to the draft scaffolds were analyzed by HiRise to estimate the genomic distance between read pairs, and the model was used to identify and break putative misjoins, score prospective joins, and make joins above a threshold. After aligning and scaffolding Chicago data,

Dovetail HiC library sequences were aligned and scaffolded following the same method. After scaffolding using Chicago and HiC library data, visual inspection of the contact maps identified two misjoin events in the largest scaffold and one other scaffold (black circles in Figure 2 - S6). Manual corrections were performed using link density plots in Juicebox (Juicebox, RRID:SCR_021172) (Robinson et al., 2018) to make breaks within those two scaffolds and produce the link density plot for the final assembly (Figure 2 - S7).

mRNA-seq library preparation

RNA was extracted separately from young (a few weeks) and old (a year) tissue (leaf and stem) from one individual using the Qiagen RNeasy kit. RNA was pooled and a library was synthesised and sequenced by Genewiz™ on an Illumina Novaseq 6000 platform in 2x150bp mode, resulting in 64,756,621 reads.

Transcriptome assembly

Raw RNA-seq reads were first cleaned by trimming the adapters using Trimmomatic v. 0.38 (Trimmomatic, RRID:SCR_011848) with the parameter "ILLUMINACLIP:TruSeq3- PE.fa:2:30:10:2:keepBothReads LEADING:3 TRAILING:3 MINLEN:36" (Bolger et al., 2014). The trimmed reads were used to assemble the transcriptome using Trinity v.2.8.5 (Trinity, RRID:SCR_013048) (Grabherr et al., 2011).

Annotation of repetitive sequences

A custom repeat library was constructed following recommendations of the MAKERP pipeline for advanced repeat construction (Campbell et al., 2014). Both structure-based and homology-based approaches were used to increase the power to detect repeats. Sequences of miniature inverted repeat transposable elements (MITEs) were collected using MITE-Hunter (Mite-Hunter, RRID:SCR_020946) (Han & Wessler, 2010) using all the default parameters. Long terminal repeat retrotransposons (LTRs) were collected using

LTRharvest (LTRharvest, RRID:SCR_018970) and LTR-digest (Ellinghaus et al., 2008; Steinbiss et al., 2009). The candidates were filtered for false positives caused by other tandem repeats such as centromeres, tandem gene duplications or other transposable elements by identifying those sequences whose alignments extend beyond the LTR boundary. Representative sequences (exemplars) were chosen as described previously (Schnable et al., 2009) to reduce the redundancy of the LTR. Then other repetitive elements were collected by first masking the genome sequence with the previously obtained MITE and LTR sequences. Then unmasked sequences were extracted and processed by RepeatModeler v.2.0.3 (RepeatModeler, RRID:SCR_015027) (Smit & Hubley, 2008) to identify additional repeats. As many repeats carry gene fragments, all the collected repetitive elements were searched against a plant protein database that contains those from swissprot plant protein and NCBI (National Center for Biotechnology Information) Refseq plants (Plant Protein Database) with transposon proteins excluded. Elements with significant hits to genes were removed along with 50bp upstream and downstream of the hit. If the remaining sequence was less than 50bp then it was completely excluded. Sequences matching the plant proteins as well as 50bp of flanking sequences were removed using the package ProtExcluder (Campbell et al., 2014). After this if the remaining portion of the sequence was shorter than 50bp, the entire sequence was excluded. Sequences of all the identified repetitive elements were joined together to form a final custom repeat library to be used to mask the repetitive elements of the genome in the Maker genome annotation protocol (Cantarel et al., 2008).To identify the type of repeat (including the repeat family), the unidentified repeats from RepeatModeler, as well as the LTRs and MITES from the custom library were searched against two transposase databases. The first was Tpases020812 (All Tansposase Protein Database). This database is composed of transposase protein sequences from the RepeatMasker v.4.1.1 (RepeatMasker, RRID:SCR_012954) (Smit et al., 2015), and from two other sources (Jiang et al., 2003; Kennedy et al., 2011), and was searched using BLASTX (BLASTX, RRID:SCR_001653) (Altschul et al., 1990). The second was the publicly available Dfam-curated library of repeats (Storer et al., 2021), which was searched using Hmmer v.3.3.1 (Hmmer, RRID:SCR_005305) (Eddy, 2009) implemented through (Dfamscan.pl). The sequences that matched the database were classified according to their top hits. Using this custom repeat library, RepeatMasker version 4.1.1 (RepeatMasker, RRID:SCR_012954) (Smit et al., 2015) was used to mask the genome, and identify the distribution of repeat types.

Genome annotation

MAKER v.3.01.03 (MAKER, RRID:SCR_005309) (Cantarel et al., 2008) genome annotation pipeline was used to annotate the genome. The input files provided for the first run were the genome assembly fasta file (Genbank accession JALGXP000000000), the reference transcriptome assembly fasta file (obtained from Trinity), and the protein homology evidence from a plant protein database (Plant Protein Database) which combines the Swissprot plant protein database and NCBI Refseq for plants excluding transposable elements. Repetitive regions were masked using our custom repeat library. Additional regions with low complexity were soft masked using RepeatMasker v.4.1.1 (RepeatMasker, RRID:SCR_012954) (Smit et al., 2015). Iterative runs of MAKER v.3.01.03 (MAKER, RRID:SCR_005309) (Cantarel et al., 2008) were undertaken in order to train the gene predictors SNAP v.2013-11-29 (Korf, 2004) and AUGUSTUS v.3.3.3 (AUGUSTUS, RRID:SCR_008417) (Stanke et al., 2006) as recommended by (Cantarel et al., 2008). The first round of annotation was based on alignments of the transcriptome to the genome. For the first round the est2genome option in the Maker control file was set to 1 to allow Maker to infer gene models directly from the RNA-seq evidence in the transcriptome. After the completion of the first round of annotations, gene models with an AED (Annotation Edit Distance) score of 0.25 or greater and a length of 50 or more amino acids were retained and used to train SNAP v.2013-11-29 (Korf, 2004) to obtain a SNAP hmm file. We then trained AUGUSTUS v.3.3.3 (AUGUSTUS, RRID:SCR_008417) (Stanke et al., 2006) using BUSCO v.3.0.2 (BUSCO, RRID:SCR_015008) (Seppey et al., 2019). First, training sequences were identified using the gene models predicted by Maker from the first run by excising regions with mRNA annotations and 1000 bp on either side. These were used to run BUSCO using the embryophyte set of conserved genes and an initial hmm model from rice. After training both SNAP and Augustus, Maker was run again, with SNAP hmm and Augustus files. A total of three rounds of training for each gene predictor were run. We used the script genestats (Card DC. Genestats) to calculate the numbers and lengths of genes, exons, introns, and UTR (untranslated region) sequences present in the predicted gene models by the final Maker run (Table 2 - S1). We ran BUSCO v.5.1.3 (BUSCO, RRID:SCR_015008) (Manni et al., 2021) with the eukaryota_odb10 lineage data set on the predicted transcript fasta file by Maker to assess the quality and the completeness of the annotated genome.

Subgenome and homeologous exchange identification

First, we identified the 20 largest scaffolds (> 40Mb; representing 1.11 Gb of the 1.22 Gb genome). These scaffolds were then aligned against themselves using Minimap2 v.2.1.8 (Minimap2, RRID:SCR_018550) (Li, 2018) to identify scaffolds that shared homology and synteny that would indicate putative homeologous chromosomes. The alignments were plotted using the R\pafr package v.0.0.2 (Winter et al., 2020) . Of these 20 scaffolds, ten pairs (with more than 50% matching across both scaffolds) were identified as the pairs of homeologous scaffolds (Figure 2 - S1). As there is no genomic data from any close diploid relatives that do not share the most recent allopolyploidization event, we clustered the putative homeologous chromosomes based on repeat abundance using kmer distributions. We partitioned the *B. decipiens* genome into subgenomes A and B by modifying the methods described in (Mitros et al., 2020). Specifically, we first identified 13 base pair sequences (13- mers) using Jellyfish v. 2.3.0 (Jellyfish, RRID:SCR_005491) (Marçais & Kingsford, 2011) and retained kmers at high abundance in the assembly (100 x or above). For each pair of scaffolds, we compared the counts of these 13-mers, identifying those that differed in abundance by three-fold or more between scaffolds. To control for any differences in scaffold length impacting this assessment, we further reduced the set of diagnostic 13-mers to those that retained a three-fold difference after standardising kmer count for the scaffold length, while keeping only those diverging in the same direction as the absolute kmer count. Hierarchical clustering of scaffolds based on difference in 13-mer counts was used to identify putative subgenomes as implemented in the R\ComplexHeatmaps package (ComplexHeatmaps, RRID:SCR_017270) (Gu et al., 2016).

We tested for homeologous exchange among the subgenomes using a Hidden Markov Model implemented in the R\HMM package (Himmelmann, 2022) . We used the most common kmer type (A or B) in 1Mbp windows as the observed states and the subgenome type for each of the 1,121 windows. The initial HMM used equal starting probabilities and transition probabilities of 0.01. We trained the HMM emission probabilities (viterbiTraining) using Scaffold 5 and Scaffold 15 as they appeared not to be subject to any subgenome exchange based on the A and B kmer density plots (Figure 2 - S3).

We examined subgenome-enriched LTRs, as differences in LTR activity in parental species can help differentiate subgenomes and can be used to assess the timing of allopolyploidy. LTRs were used for this because they are rapidly evolving, making it easy to differentiate between related sub families. The timing of insertions can be calculated by examining the substitution rates for members of the same subfamily using the 5' and 3' regions (Mitros et al., 2020). Specifically, intact retrotransposons in the genome were identified using LTR HARVEST (Ellinghaus et al., 2008). The 'best' option was used for pairing overlapping LTR sequences, allowing the inner sequences of the retrotransposons to contain gaps. We performed an all-versus-all BLAST (BLAST, RRID:SCR_004870) (Camacho et al., 2009) on the long terminal repeat segments of the identified LTRs with an e-value cut off of 1e-2. Hits with the percentage of alignment between query and subject equal or greater than 90% over their entire length were selected. We then used MCL algorithm (van Dongen & Abreu Goodger, 2012) to cluster the filtered blast alignments into retrotransposon subfamilies using an inflation parameter (-I) of 3. We counted the occurrence and the total base pairs that each LTR subfamily obtained from above clustering in the putative A and B subgenomes identified above. We identified LTR subfamilies that were three times more common in one of the subgenomes using both occurrence and bp count. Then we determined if these repeats overlapped with multiple A or B genome preferred kmers to confirm that kmers were representing longer repetitive sequences and to confirm that these kmers were marking repeat expansion that occurred just before the allopolyploidy event.

Gene function prediction

The predicted protein sequences obtained from the final run of MAKER were aligned to the UniProtKB/Swiss-Prot (UniProt Consortium, 2021) and TAIR10 (Berardini et al., 2015)
protein databases using BLASTP (BLASTP, RRID:SCR_001010) (Camacho et al., 2009) with an e-value cutoff of 1.0e-5. The GO term associated with the best hit for each BLASTP search was identified in each of the three databases above and assigned to the *B. decipiens* query. InterProScan v. 5.51-85.0 (InterProScan, RRID:SCR_005829) (Quevillon et al., 2005) was used to search the query protein fasta against the Pfam (Punta et al., 2012) protein family database and identify functional protein domains. Pfam

accessions and GO terms were retrieved for the *B. decipiens* query sequences. The query protein sequences were BLAST (Basic Local Alignment Search Tool) searched against the KEGG database (Kanehisa, 2019; Kanehisa et al., 2021; Kanehisa & Goto, 2000) using the online tool KofamKOALA - KEGG orthology search (Aramaki et al., 2020) with an e-value cutoff of 1.0e-5. KEGG orthology terms and enzyme codes were retrieved for each hit.

Genome synteny and whole genome duplication

To determine if the *B. decipiens* genome had undergone whole genome duplication, a reciprocal BLASTP (BLASTP, RRID:SCR_001010) was conducted using *B. decipiens* protein sequences as the query against themselves with a minimum e-value greater than le-5. Then MCScanX (Wang et al., 2012) was used to identify syntenic blocks within the genome. The collinear blocks obtained via MCScanx between the putative A and the B subgenome of *B. decipiens* were visualised using Synvisio (Bandi & Gutwin, 2020). Similarly, we identified collinear blocks between the two putative subgenomes of *B. decipiens* and *S. bicolor* by conducting a reciprocal BLASTP (BLASTP, RRID:SCR_001010) comparing protein sequences from each species using MCScanX and plotted the results using Synvisio.

Estimating the timeline of paleo tetraploidy

We estimated the timing of speciation events in the Andropogoneae using *Panicum hallii* and *Setaria italica* as outgroups. The reference gene sets for *Sorghum bicolor* v3.1.1, *Panicum hallii* v2.2, *Setaria italica* v2.1, *M. sinensis* v7.1 and *Zea mays* (B73 RefGen_v4) were downloaded from Phytozome v12.1 (Phytozome, RRID:SCR_006507) (Goodstein et al., 2012). The *Saccharum spontaneum* reference gene set (Zhang et al., 2018) was also downloaded. We separated the A and B subgenomes of *M. sinensis* (as identified in (Mitros et al., 2020)), as well as those of *B. decipiens*, to compare the timeline of the paleo allopolyploidy events between *Miscanthus* and *B. decipiens* species. We identified 1:1 orthologs between all species (or subgenomes) using OrthoFinder v.2.3.8 (OrthoFinder, RRID:SCR_017118) (Emms & Kelly, 2019). A core set of 392

single-copy genes were retained and multiple sequence alignments were performed for each orthologous cluster using Genodup (Mao, 2019) and Mafft (Mafft, RRID:SCR_011811) (Katoh & Standley, 2013). Poorly aligned regions were removed using Gblocks 0.91b (Gblocks, RRID:SCR_015945) (Talavera & Castresana, 2007) and the final alignments were concatenated into a single alignment.

Phylogenomic analysis

The best model of evolution was inferred using jModelTest2 (Darriba et al., 2012; Guindon & Gascuel, 2003). A phylogenetic tree was constructed using RAxML (RAxML, RRID:SCR_006086) (Stamatakis, 2014) with the GTRGAMMAI model of evolution and 1000 bootstrap replicates. *Setaria italica* and *Panicum hallii* were designated as outgroups.

Divergence date estimation

Divergence among lineages in the phylogeny were estimated from the concatenated alignment using BEAST v.2.5 (BEAST, RRID:SCR_010228) (Bouckaert et al., 2019; Suchard et al., 2018) after using bModelTest (Bouckaert & Drummond, 2017) to infer the best substitution model. Parameters included the GTR (general time reversible) substitution model with unequal frequencies, four gamma categories, estimated shape and invariant sites. We chose a relaxed log normal clock with estimated clock rates. We set priors to constrain the estimated dates at the *Setaria-Panicum* (12.8-20 MYA), and the Andropogoneae (13-21.2 MYA) nodes, using a uniform distribution between the minimum age and maximum ages of divergence times obtained from the TimeTree database (Kumar et al., 2017). BEAST2 (BEAST2, RRID:SCR_017307) (Bouckaert et al., 2019) analysis was conducted for 50 million generations, and logging at every 5000 trees. Convergence between runs was assessed with Tracer v.1.6 (Rambaut et al., 2018). Trees were summarised with TreeAnnotator (Suchard et al., 2018) using a burn-in value of 20%.

Timing of subgenome-specific LTR expansions

We aligned the long terminal repeats of each LTR family cluster using Mafft (Mafft, RRID:SCR_011811) (Katoh & Standley, 2013). We computed Jukes-Cantor distance matrices using the R\ape package (Paradis & Schliep, 2019). We estimated the divergence times of each LTR family as k/2r (k=divergence, r=substitution rate) (Moniz de Sá & Drouin, 1996). We used 1.3×10^{-8} as the substitution rate per site per year (Ma & Bennetzen, 2004).

Determination of biases in subgenome gene retention

We analysed the collinear blocks between *S. bicolor* and *B. decipiens* to assess differences in retention of duplicated genes between subgenomes. Subgenome-specific retention was inferred as the number of genes retained in a given subgenome divided by the number of inferred ancestral (i.e., pre duplication) gene numbers. Consequently, we calculated the number of ancestral (pre duplication) genes as those orthologous genes present in *S. bicolor* and in one or both of the two subgenomes. We then compared this number to the total number of genes present only in subgenome A, only in subgenome B or in both subgenomes. We then used a two-sided Fisher's exact test to determine if there was a significant difference in retention of genes between the subgenomes under the null hypothesis that gene loss between subgenomes was random. We also used the results from OrthoFinder (OrthoFinder, RRID:SCR_017118) to confirm this pattern. Specifically, we identified 1:1 and 1:2 orthologs between *S. bicolor* and *B. decipiens*, only retaining 1:2 orthologs that were mapped to chromosomes on both *B. decipiens* subgenomes. Differences in the function of duplicated genes retained after the WGD compared to those returning to single copy was tested through gene enrichment analysis using R\topGo (topGO, RRID:SCR_014798) (Alexa & Rahnenfuhrer, 2006). All single-copy genes in A or B, or duplicated genes in both A and B were used as foreground genes and the remaining ancestral genes (retained duplicated or single copy) as background genes.

Data Availability

Raw reads for the genome assembly have been deposited under BioProject accession number PRJNA819081. Illumina library raw reads namely 10x, Chicago and HiC data

have been deposited in the Sequence Read Archive (SRA) under study accession numbers SRR18458736, SRR18471564 and SRR18471563. RNAseq data have been deposited under SRA accession number SRR18471562. Genome assembly of *B. decipeins* is deposited in the NCBI genome database under the accession JALGXP000000000.

Abbreviations

℃ : degree Celsius, BCD : *Bothriochloa Capillipedium* and *Dichanthium*, BLAST : Basic Local Alignment Search Tool, BLASTP :Basic Local Alignment Search Tool Program, bp : base pairs, BUSCO : Benchmarking Universal Single-Copy Orthologs, DNA: Deoxyribonucleic Acid, FCM :flow cytometry, Gb : Giga bases, gDNA :genomic DNA, GO : Gene ontology, HMM :Hidden Markov Model, KEGG : Kyoto Encyclopedia of Genes and Genomes, LTR : long terminal repeats, Mb : mega base pairs, mg :milligram, MITE : miniature inverted repeat transposable element, mL :millilitre, mM : millimolar, mm : Millimetre, mRNA : Messenger RNA, MYA : Million Years Ago, NCBI : National Centre for Biotechnology Information, ng : nanogram, NSW : New South Wales, PCR : polymerase chain reaction, RNA : ribonucleic acid, Seq :sequencing, TE : transposable elements, UTR : untranslated region, WGD : Whole genome duplication, µg : microgram, µL : microliters, µm : micrometre

Competing Interests

The authors declare that they have no competing interests.

Funding

This study was supported by Hermon Slade Foundation - Grant number HSF1703, Monash Graduate Scholarship, Monash University, Monash International Tuition

Scholarship, Monash University and Denis and Maisie Carr Award and Travel grant 2020, School of Biological Sciences, Monash University.

Author's Contributions

K. A. H. - conceived and designed the study. K.A.H., J.L.M. and A.F.L. - supervised the research, C.L. conducted the FCM, RNA extractions and phylogenetic analyses, K.A.H. and N.P.D. - annotated the genome, carried out the subgenome and homeologous exchange identification, genome synteny and whole genome duplication analysis and timing of subgenome-specific LTR expansion, N.P.D. carried out repeat library construction, transcriptome assembly, N.P.D. and P.B. carried out functional predictions, determination of biases in subgenome gene retention. N.P.D and K.A.H. wrote the first draft of the manuscript. All authors approved the final version of the manuscript.

Acknowledgements

We thank Brendon Levot for collecting the seeds from which the individual specimen used in this study was grown from. Extraction of genomic DNA, preparation and sequencing of libraries and genome assembly was done by Dovetail Genomics. The preparation and sequencing of RNA-seq libraries was done by Genewiz. Permission to collect plant material for the study was granted by NSW National Park and Wildlife Service - Licence number- SL 102026.

Appendix I - Supplemental information for : Genome assembly of an Australian native grass species reveals a recent whole genome duplication and biased gene retention of genes involved in stress response

Table 2 - S1. Summary of the annotation for protein coding genes.

Primary transcripts (i.e., longest at each locus)	60,652
Average exon number/gene	5.46
Median transcript length (bp)	2880
Median CDS length (bp)	1017
Median exon length (bp)	269
Median intron length (bp)	359
Median 3’ UTR length (bp)	292
Median 5' UTR length (bp)	135

Table 2 - S2. Length of genome sequence repeat-masked by repeat element class.

Class	**Mbp**	**% Of genome**

Gypsy	237.21	19.47
Copia	105.38	8.65
DNA	40.64	3.34
LINEs	25.61	2.10
Simple	6.73	0.55
Low complexity	0.93	0.08

Table 2 - S3 Summary of the first 10 most significant GO terms among overrepresented genes in genes retained as single copies as reported by the gene enrichment analysis.

GO ID	**Term**	**Annotated**	**Significant**	**Expected**	**P-value**

GO:0009451	RNA modification	459	128	74	3.6×10^{-12}
GO:0120009	Intermembrane lipid transfer	14	11	2.26	3.3×10^{-7}

GO:0035627	Ceramide transport	14	11	2.26	3.3×10^{-7}
GO:0009846	Pollen germination	136	39	21.93	9.1×10^{-5}
GO:0071585	Detoxification of cadmium ion	14	7	2.26	9.6×10^{-5}
GO:0009658	Chloroplast organization†	442	92	71.26	2.1×10^{-4}
GO:0006397	mRNA processing	943	190	152.04	2.8×10^{-4}

GO:0071712	ER-associated misfolded protein catabolism†	23	11	3.71	3.0×10^{-4}
GO:1900865	Chloroplast RNA modification†	53	22	8.55	3.3×10^{-4}
GO:0010821	Regulation of mitochondrion organization†	17	10	2.74	4.5×10^{-4}

†Terms associated with organelle functions

Table 2 - S4 Summary of the first 10 most significant GO terms among overrepresented genes in genes retained as duplicates as reported by the gene enrichment analysis.

GO ID	Term	Annotated	Significant	Expected	P-value
GO:0045893	Positive regulation of transcription†	962	550	455.35	7.4×10^{-18}
GO:0045489	Pectin biosynthetic process	147	109	69.58	1.2×10^{-11}
GO:0045892	Negative regulation of transcription†	593	333	280.69	7.0×10^{-9}
GO:0006355	Regulation of transcription†	3176	1759	1503.31	1.9×10^{-8}
GO:0009958	Positive gravitropism	48	41	22.72	1.9×10^{-8}
GO:0006357	Regulation of transcription by RNA polymerase†	836	456	395.71	6.5×10^{-8}
GO:0035556	Intracellular signal transduction	952	494	450.61	2.2×10^{-7}

GO:0010222	Stem vascular tissue pattern formation	18	18	8.52	1.9×10^{-7}
GO:0015031	Protein transport	1674	910	792.36	1.6×10^{-6}
GO:0009414	Response to water deprivation†	963	460	455.82	1.7×10^{-6}

†Terms associated with transcription and responding to external stress

Table 2 - S5. Illumina shotgun libraries for *Bothriochloa decipiens*.

Library Type	Bases (billions)	Read pairs (millions)	Read length
10X	120	400 million	2x150bp
Chicago	140	460 million	2x150bp
HiC	120	400 million	2x150bp

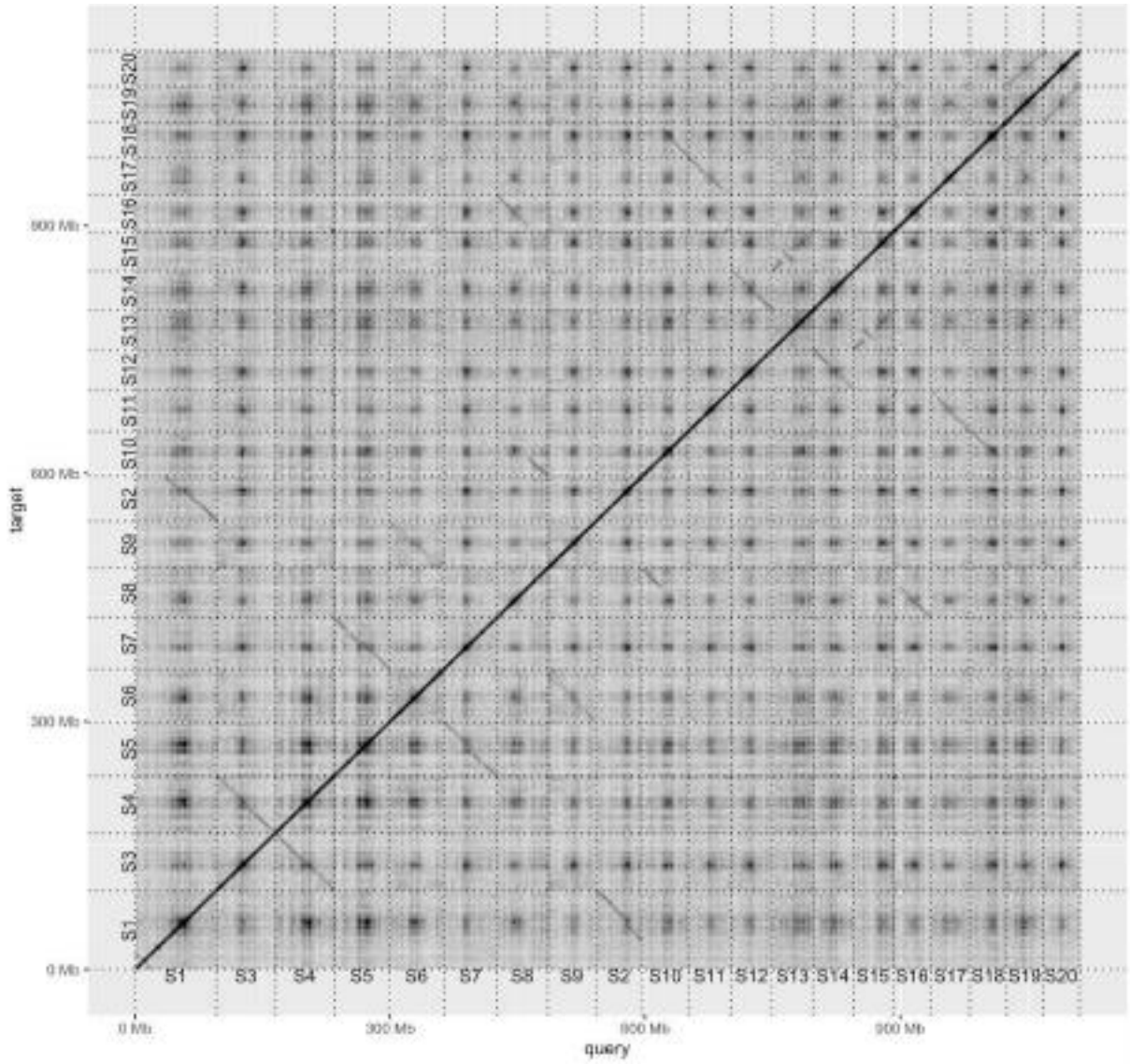


Figure 2 - S1 Mapping of the *Bothriochloa decipiens* genome against itself to identify putative homeologous chromosomes.

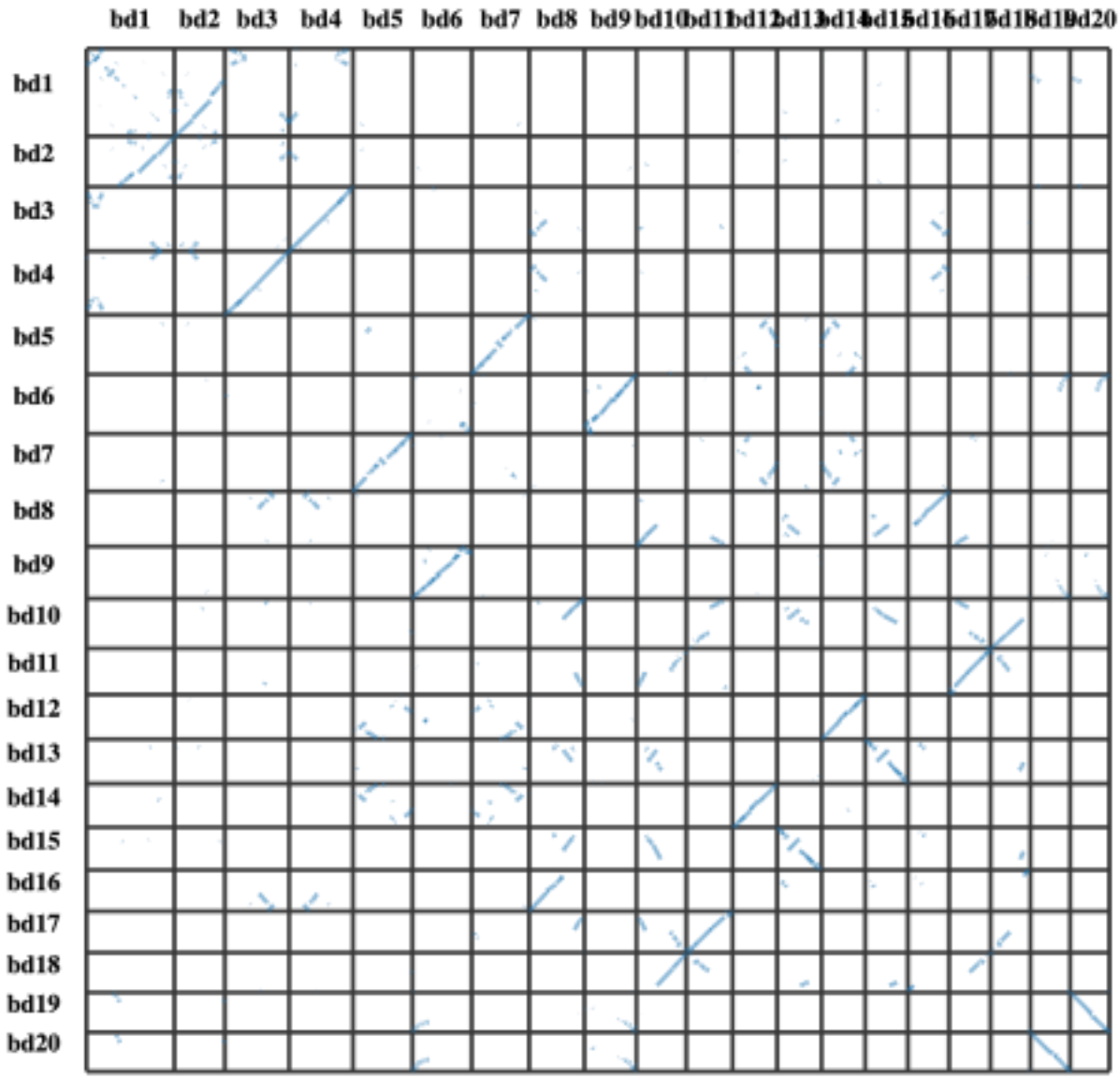


Figure 2 - S2 The dot plot depicts the syntenic relationship of the 20 chromosomes of *B. decipiens* (bd) aligned against themselves.

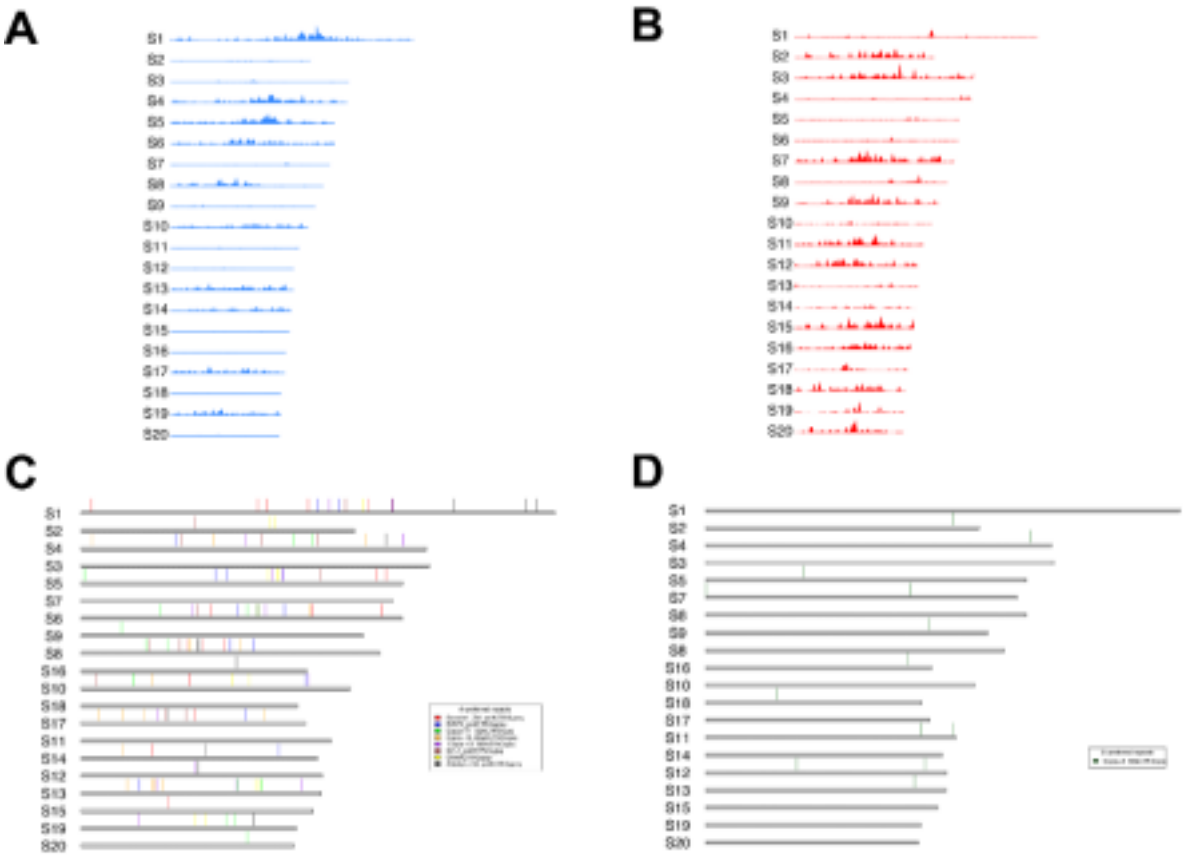


Figure 2 - S3 (A) Density of subgenome A preferred kmers. (B) Density of subgenome B preferred kmers. (C) Repeats marked by overrepresented kmers in subgenome A. (D) Repeats marked by overrepresented kmers in subgenome B. The scaffolds belonging to putative subgenomes (Figure 2 - 3) are paired together as homeologs one after another. These likely reflect repeat expansions during the period the species involved in the paleopolyploid event were separated.

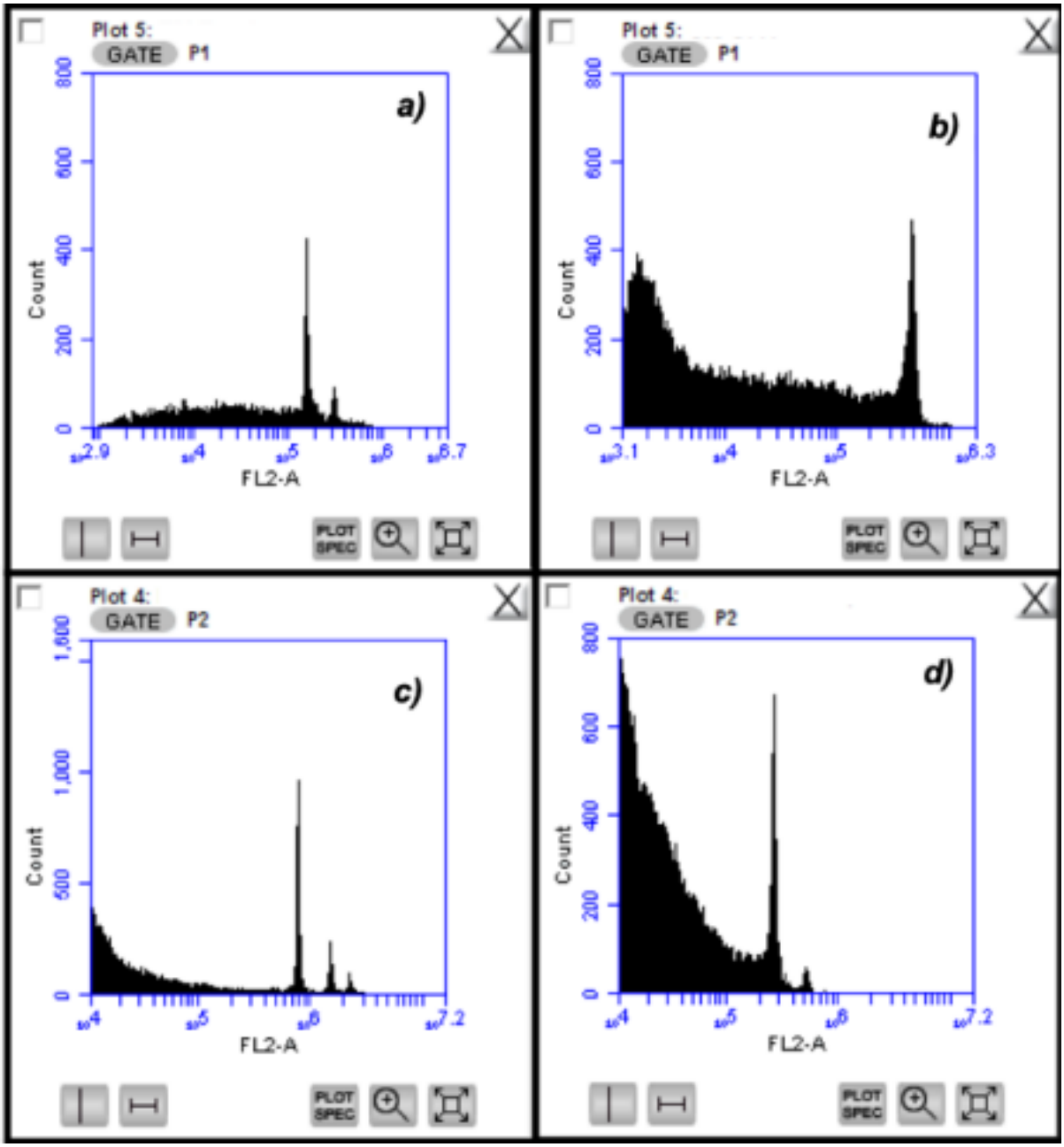


Figure 2 - S4. Indicative flow cytometry results of *Solanum* and *Pisum* size standards, and 2C genome sizes of diploid *Bothriochloa decipiens* and tetraploid *Bothriochloa macra*. *(a) Solanum*; FL2 mean = 182,944. *(b) B. decipiens*; FL2 mean = 249,158. (*c*) *Pisum*; FL2 mean

= 777095. (*d*) *B. macra*; FL2 mean =480804. Smaller, additional peaks represent the smaller proportion of cells in G2 phase of cell division.

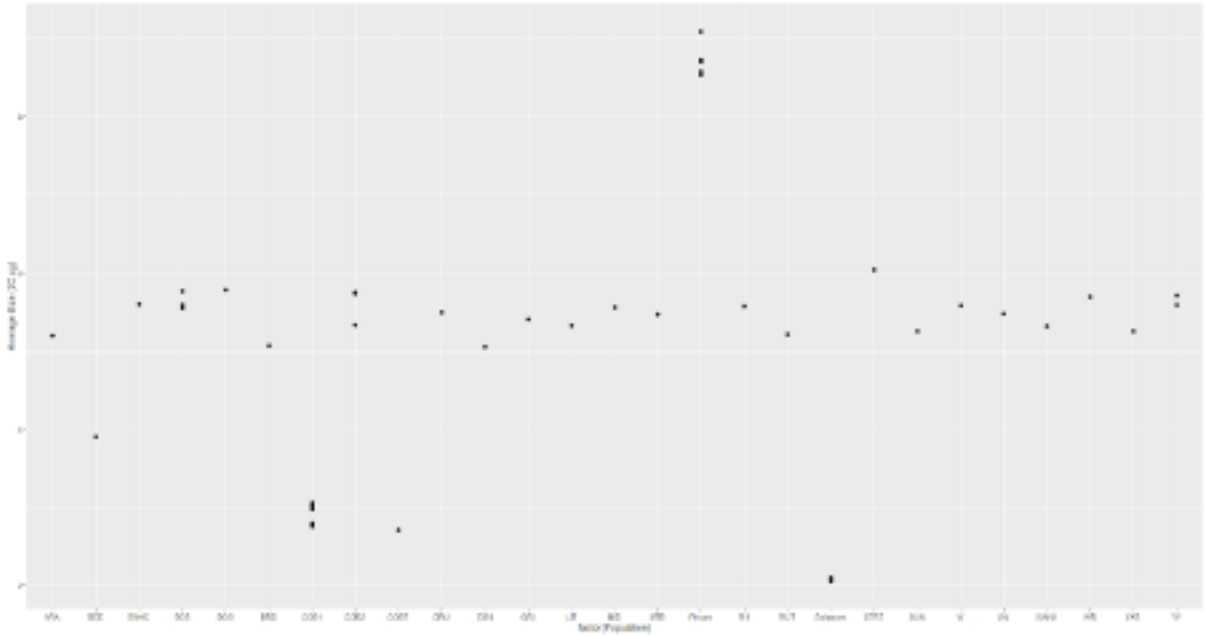


Figure 2 - S5. Estimated genome size for samples from all populations, and two internal standards (*Pisum sativum* and *Solanum lycopersicum*).

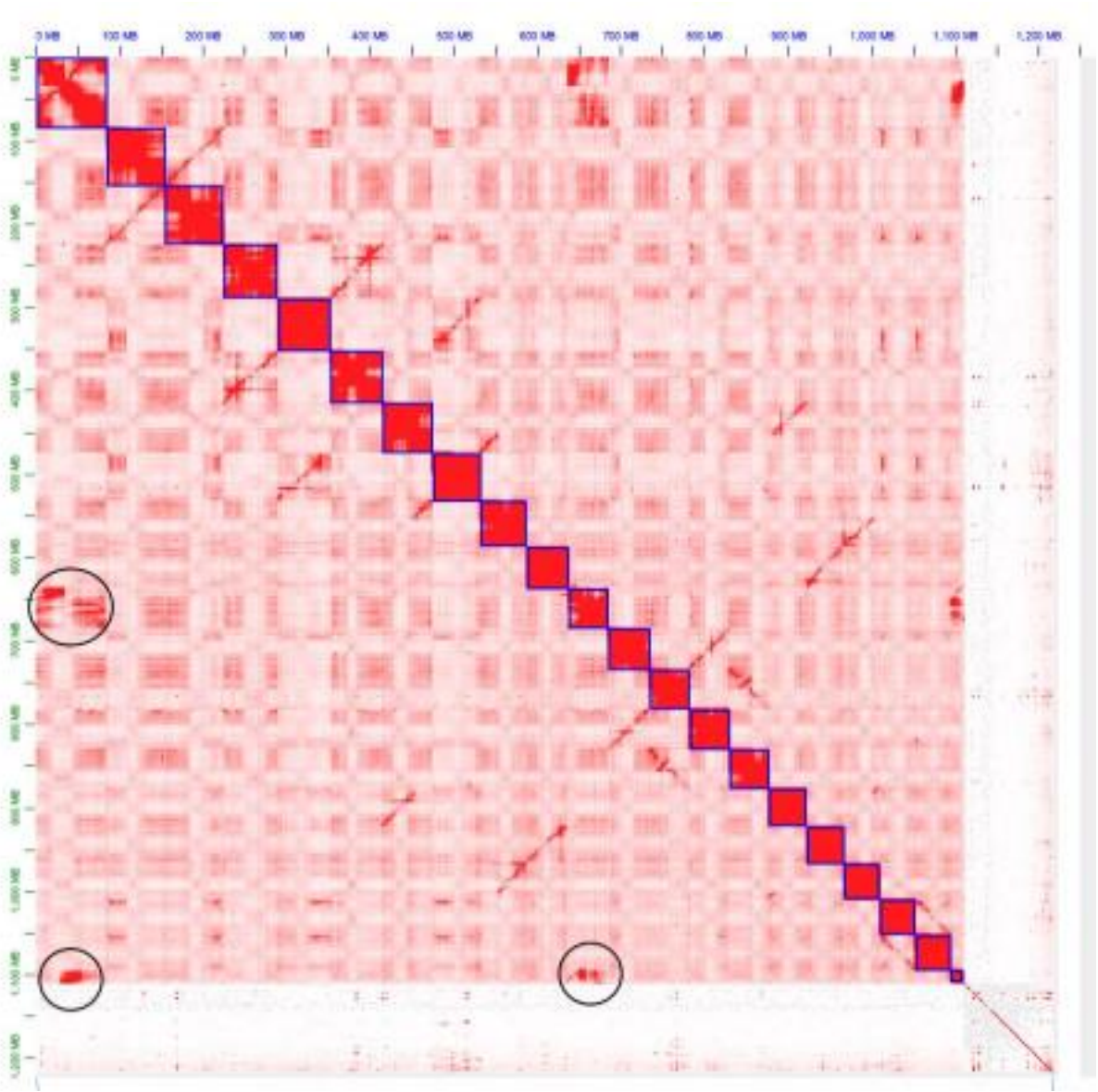


Figure 2 - S6. Mis joins in two scaffolds of the first assembly before the manual edits are shown in black circles.

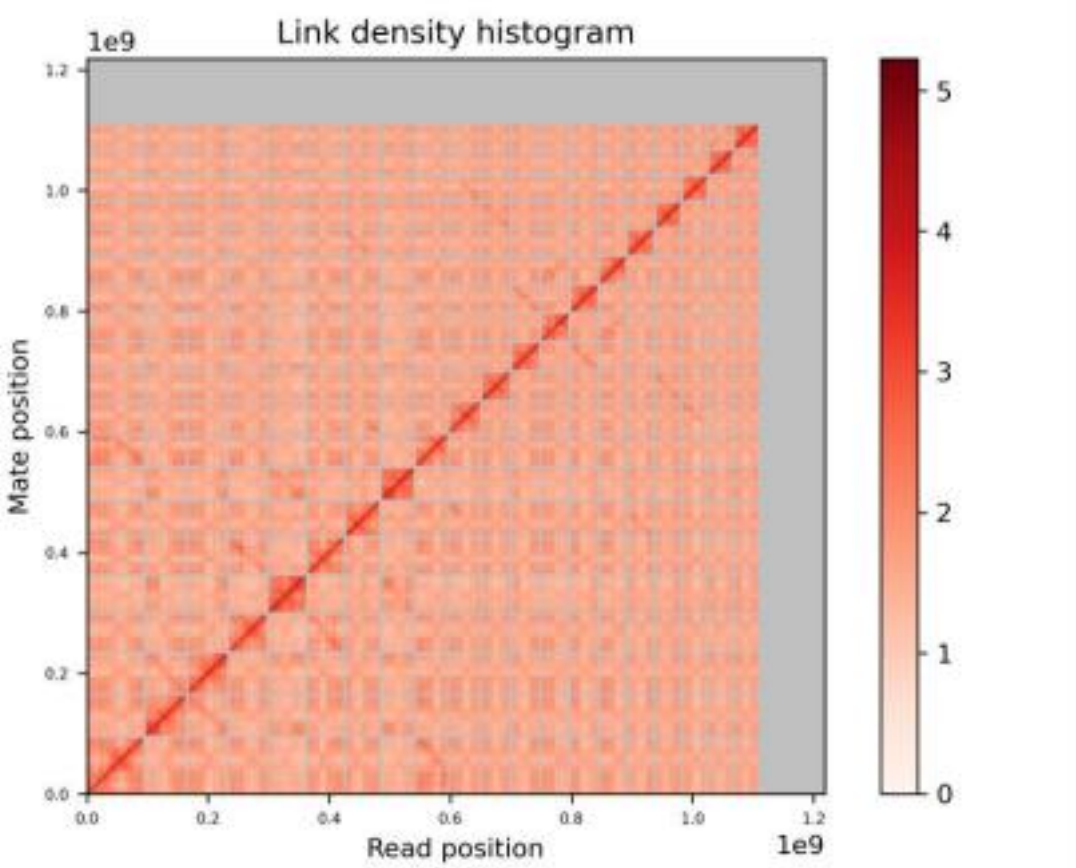


Figure 2 - S7. Scaffold Hi-C contact map. The x and y axes give the mapping positions of the first and second read in the read pair respectively, grouped into bins. The colour of each square gives the number of read pairs within that bin. White vertical and black horizontal lines have been added to show the borders between scaffolds. Scaffolds less than 1 Mb are excluded

Chapter 3 - Signals of climate adaptation and predictions of future maladaptation in an Australian native grass

Nissanka P. De Silva[1*], Christopher Lee[1], Paul Battlay[1], A. Fournier-Level[2], Joslin L. Moore[3] , Kathryn A. Hodgins[1]

[1]School of Biological Sciences, Monash University, Clayton, Victoria, Australia, 3800. [2]School of BioSciences, University of Melbourne, Melbourne, Victoria, Australia, 3010. [3]Arthur Rylah Institute for Environment Research, Heidelberg, Victoria, Australia, 3084.

[*]Corresponding author

In preparation for *Evolutionary Applications*

Abstract

Australian grasslands have been degraded due to anthropogenic land use change while the threat of further degradation increases with predicted climate change. Australian native grass species *Bothriochloa decipiens* is important for grassland rehabilitation due to its ability to establish well from direct seeding on many soil types, and the ability to withstand pressure caused by overgrazing. Studies that investigate the genetic potential of Australian grasses and their suitability to be used in climate change resilient restoration is limited in literature. Using Genotype by Sequencing, we identified 226,197 SNPs in 80 individuals and used these data to examine population structure as well as signals of climate adaptation across the genome of the *B. decipiens*. We discovered evidence of significant population differentiation (mean F_{st} = 0.42) that was structured geographically suggesting local pollen and seed dispersal. Using genotype-environment associations, we discovered evidence of climate adaptation in *Bothriochloa decipiens*. Loci related to drought stress response were overrepresented, suggesting that aridity may be important in structuring adaptive genetic variation in this species. Genomic offset analysis identified populations in southwestern regions of the range as more vulnerable to maladaptation under climate change. This study provides the critical first step in assessing the suitability of using *B. decipiens* for climate change resilient restoration through assisted gene flow. Given the genomic signals of local climate adaptation combined with local dispersal, *B. decipiens* may benefit from assisted gene flow under climate change.

Keywords - assisted gene flow, climate adaptation, population genomics, landscape genomics, genomic offset

Introduction

Rapid climate change is predicted to be a key driver of biodiversity decline and/or loss in the next century (Dawson et al., 2011; Pacifici et al., 2015; Warren et al., 2013) and widespread climate induced local extinctions have already been reported (Urban, 2015; Wiens, 2016). In response to changing climates species are predicted to either migrate or adapt *in situ* through plasticity or evolutionary adaptation (Hoffmann et al., 2021). There is abundant evidence that both plasticity and adaptation have already been key in many species' responses to climate change (Scheffers et al., 2016). But with climatic extremes, the buffering that adaptive plasticity provides is expected to reach its limit, meaning that evolutionary change will be critical.

Efforts to conserve native ecosystems through restoration activities are increasing worldwide (Cunningham, 2008; Stange et al., 2021). However, as such restoration is carried out in an era of accelerated climate change, an assessment of the extent of climate adaptation and predictions regarding the extent of future climate maladaptation, particularly for key species that are frequently the backbone of restoration, could help direct mitigation efforts such as assisted gene flow (AGF). Assisted gene flow facilitates *in situ* adaptation to climate change via human mediated movement of individuals and/or their gametes within the species range (Aitken & Bemmels, 2016; Aitken & Whitlock, 2013; Broadhurst et al., 2008; Weeks et al., 2011). Existing populations can benefit via assisted gene flow by introduction of pre-adapted genotypes to new local climates or by increasing the frequency of these genotypes (Broadhurst et al., 2008; Kreyling et al., 2011; Weeks et al., 2011).

Although the main focus of practising assisted gene flow in restoration is speeding climate adaptation (Aitken & Whitlock, 2013), there are risks associated with introducing genotypes from foreign populations to candidate recipient populations. If ploidy differences and large-scale chromosomal differences between populations are present, assisted gene flow can cause the decline of the resultant population due to outbreeding depression (Edmands, 2007) by causing partial or complete sterility (Frankham et al. 2011). Genetic integrity of the population could be lost and potential genomic extinction

can occur during hybridization between distinct populations (Todesco et al. 2016). Alleles carried over by foreign populations can be maladaptive to the local conditions or the combination of them with local alleles may cause outbreeding depression (Edmands 2007). Investigating population genetic structure is important to ensure the potential seed transfer between populations with similar genetic structure and variability (Broadhurst et al., 2008), as significant differentiation of populations, may signal the potential for such incompatibilities to arise. For a successful ecological restoration process based on assisted gene flow, knowledge of both the population genetic structure and variation of local adaptation across the landscape is therefore essential (Hufford & Mazer, 2003; Knapp & Rice, 1994; Weeks et al., 2011).

Common garden and reciprocal transplant experiments are classic approaches for assessing the genetic basis of phenotypic variation and local adaptation to traits associated with climate (Baughman et al., 2019; Clausen et al., 1940; De Kort et al., 2013; Hereford, 2009; Langlet, 1971; Leimu & Fischer, 2008; Savolainen et al., 2007). However with the recent technological advancements in high throughput genome wide DNA sequencing (Allendorf et al. 2010; Morin et al. 2004; Narum and Hess 2011), the use of genetic markers in natural populations to detect local adaptation in non-model species has increased (Luikart et al., 2003). Environmental and population genetic data can be used to detect changes in allele frequency associated with spatial environmental variation through genotype-environment association (GEA) analyses (Forester et al., 2018; Günther & Coop, 2013; Jordan et al., 2017; Rellstab et al., 2015). These approaches to identify the genetic basis of adaptation have the ability to control for confounding effects due to spatial genomic structure (e.g., IBD) and help explain the basis of population divergence caused by more restoration relevant environmental variation (Bradburd et al., 2013; Lotterhos & Whitlock, 2015; Wang & Bradburd, 2014). GEA analyses are relatively effective at detecting weak, polygenic signals of selection while displaying robustness to confounding factors (Capblancq et al., 2018; Forester et al., 2016; Lotterhos & Whitlock, 2015). GEA analyses are increasingly used to detect signals of local adaptation associated with environmental drivers related to them with implications for restoration (Brauer et al., 2016; Perrier et al., 2018) and also to predict population responses to future climate change and the genetic vulnerability experienced by populations with environmental change (Bay et al., 2018; Fitzpatrick & Keller, 2015; Jia et al., 2020; Lu et al., 2019;

Varas-Myrik et al., 2022). Such information could therefore be used to inform management practices to mitigate the effects of climate change through assisted gene flow.

Grassland communities in Australia are under significant threat (Hobbs & Yates, 2000). In south- eastern Australia, less than 1% of native grasslands remain, in dwindling remnant patches (Lunt & Morgan, 2002). Further, many areas in Australia that support grasslands are predicted to be strongly impacted by climate change (Timbal, 2009). Consequently, conserving genetic diversity of grasses, especially adaptive genetic diversity will be important for maintaining evolutionary resilience of populations (Dunlop et al., 2012). Research focussing on genetic diversity in grasses is mainly centred around species of economic importance (Buckler et al., 2001), and there are few studies that have examined the genetic diversity of ecologically important native grass species.

Bothriochloa decipiens (blue pitted grass) is a warm season, perennial, tufted grass that is widespread in subtropical NSW and Queensland as well as tropical Queensland (Simon & Alfonso, 2011). Due to its ability to establish well from direct seeding on many soil types, and the ability to withstand pressure caused by overgrazing, it has become an important species for grassland rehabilitation (Simon & Alfonso, 2011). It is a diploidized allotetraploid species that belongs to a group with a complex history of hybridization and polyploidy (Chapter 2). Indeed, many grasses, including *B. decipiens*, have experienced relatively recent whole genome duplication events. The gene expression changes and phenotypic alterations that follow gene duplication via polyploidy is a significant source of evolutionary novelty in plants (Flagel & Wendel, 2009). Some have hypothesised that allopolyploidy has given grasses the ability to adapt to new environments, thereby enabling successful expansion and establishment (Godfree et al., 2017; Linder & Barker, 2014). Evidence supportive of this hypothesis was reported in *B. decipiens* where genes related to stress response were retained as duplicated following a recent whole-genome duplication (Chapter 2). Hence, we hypothesise that retained duplicated genes will also be a key contributor to local adaptation to climate within the species.

Our goal was to examine the adaptive genetic landscape of the native grass species *B.*

decipiens. To do this we conducted a landscape genomic study of this species using eight (8- 18 individuals per population) populations and 226,197 SNP markers identified using genotype by sequencing (GBS). Specifically, we addressed the following specific questions. 1) What are the patterns of population genetic differentiation and diversity across the landscape? Understanding population structure and diversity could potentially provide information on regions that are genetically depauperate and potentially vulnerable to environmental change as well as regions that are highly divergent and more prone to outbreeding depression when mixed. 2) Are there regions of the genome that show signatures of climate adaptation? The extent of local adaptation to climate in this long-lived perennial grass has not been investigated despite its potential importance for grassland restoration. 3) What are the features of the climate adaptation candidates? We predict that regions of the genome that show signatures of climate adaptation will be enriched for stress tolerance genes and duplicated genes. 4) What populations are most at risk of maladaptation under climate change scenarios? Geographic regions most at risk could be considered as candidates for future assisted gene flow during restoration.

Methods

Sample collection and visualising the species distribution across aridity gradients

Leaf samples of 88 individuals belonging to 8 populations across NSW and Queensland were collected during April 2021. As *B. decipiens* is morphologically similar to *B. macra*, during collection special attention was paid in observing the anthers of the flowers of the sampling individuals. Single anther of *B. decipiens* sets it apart morphologically from *B. macra* which has three anthers (Sumadijaya, 2015). Therefore, populations with individuals bearing flowers with a single anther were identified for collection. The locations of the populations are given in the map below (Figure 3 -1). Sampling locations were selected with at least 10 km distance between each to ensure population differentiation. From each population leaf samples of about 15 individuals were collected and were placed in envelopes with silica beads, to dry the leaves upon collection. After collection, ploidy of the collected populations was checked using flow cytometry (Dolezel et al., 2007) and diploid populations were selected to be used in the analyses of

this study. The abbreviations used to identify the populations of the samples used in this study, their geographic location in terms of latitude and longitude coordinates, the state which each population belongs to, and the number of samples used from each population is summarised in Table 3 - S1. To visualise how this species is distributed across aridity gradients, potential evapotranspiration (PET) measures were extracted using randomly generated geographic coordinates within the Australian continent from the database of global aridity index from the consortium for spatial information (Trabucco & Zomer, 2018) and for each point source aridity was calculated by dividing the PET by the annual rainfall. Here what is referred to as 'source aridity', is the inverse of the aridity index (Aridity Index (AI) is a simple but convenient numerical indicator of aridity based on long-term climatic water deficits and is calculated as the ratio P/PET, P=rainfall and PET= potential evapotranspiration) (Grainger 1994). Here the inverse of the aridity index was used as it is easy to interpret as larger values correspond to greater aridity. Distribution of source aridity along the geographic locations were plotted using the R\viridis package (Garnier et al., 2021) and the locations representing the species distribution were overlapped on the source aridity distribution (Figure 3 -1).

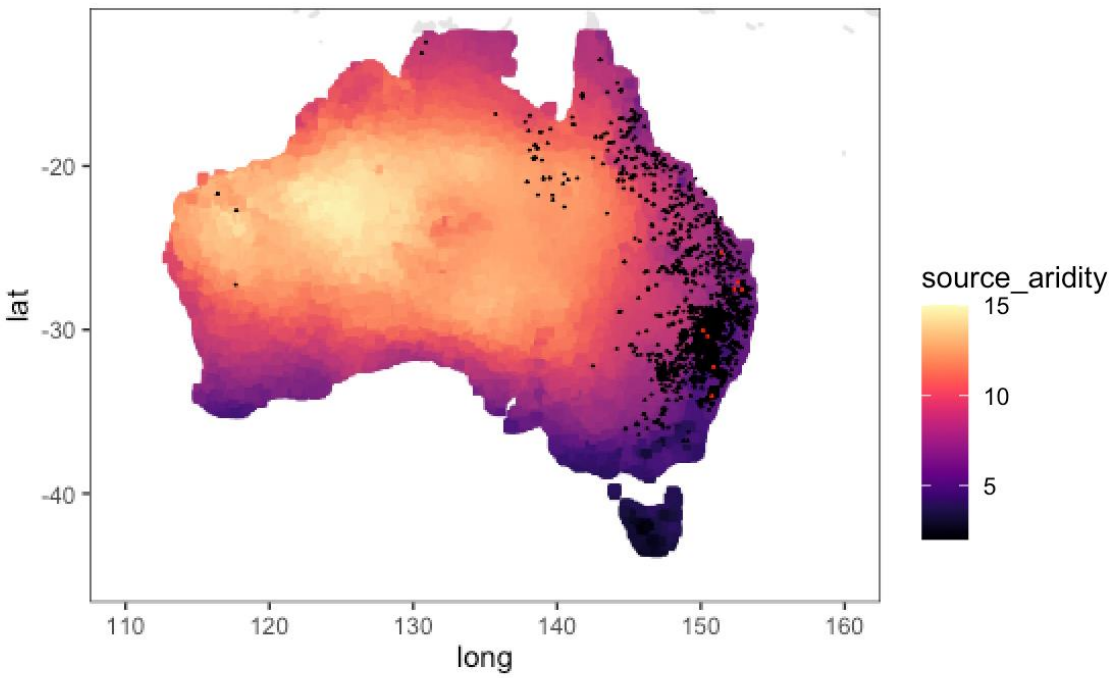


Figure 3 - 1. Distribution of *B. decipiens* across aridity gradients. Map depicting how aridity varies across the species distribution range of *B. decipiens*. Black dots represent locations of

reported occurrences of this species according to records from Atlas of Living Australia. Red dots represent the locations of the eight sampling sites of the populations used in this study.

DNA extraction, preparation of GBS libraries and sequencing

Genomic DNA was extracted using 30-40 mg of dried leaf tissues from the 88 samples using the DNAeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). The quality of the extracted DNA was checked by running samples on 2% agarose gel stained with ethidium bromide and viewing it under UV-Vis Spectrometer (NanoDrop 1000, Thermo Scientific). DNA quantity was assessed using a QuBit broad-sensitivity DNA quantification system (Invitrogen, Carlsbad, CA, USA). GBS libraries were prepared by following a modified version of the protocol from (Poland et al., 2012). This involves a double digest (with *HF-PstI* and *MspI* restriction enzymes), genotype by sequencing (ddGBS) for equal amounts of DNA per individual. To modify the protocol, PCR reactions were performed for individual samples and equal amounts of PCR products were pooled. Fragments between 400-600 bp were selected and sequencing of the libraries was done by GENEWIZ^{TM} (Suzhou, China) on the Illumina HiSeq 4000 platform (paired end, 150 bp).

SNP calling

Raw reads were demultiplexed and cleaned by filtering for low quality reads using STACKS (2.2) process_radtags (Catchen et al., 2013) with default settings. Then the reads were aligned to the *B. decipiens* reference genome using Burrows-Wheeler Aligner (BWA) (Li & Durbin, 2009). Variants were called using GATK HaplotypeCaller (Poplin et al., 2017). Variant filtering was done as follows using VCFtools (Danecek et al., 2011). First, we removed individuals with fewer than 25,000 reads as well as indels. We also removed low coverage individual genotypes (min depth 5). Hard filtering was performed using GATK using the following criteria: QualityByDepth(QD)<2.0, FisherStrand(FS)>60.0, StrandOddRatio(SOR)>3.0, RMSMappingQuality(MQ)<40, ReadPosRankSum < -8.0 and MQRankSum< -12.5. We refer to this SNP set after the above filtering steps as the *initial filtered dataset*. This *initial filtered dataset* was filtered again using VCFtools. Genotypes with genotype quality less than 20 were set as missing and sites with more than 20% missing data were excluded. Then individuals with more than 50% missing data were removed, filtered variants with a minor allele frequency less than 0.05. The resulting set of variants were filtered to retain only biallelic SNPs. Then we

removed variants which showed more than 80% of observed heterozygosity and finally thinned the data set for linkage using a single SNP per 1kb window, resulting in 44,152 SNPs from 81 individuals. This SNP set will be referred to as the *globally thinned dataset*. The *initial filtered dataset* was used in another less stringent filtering procedure as follows, using VCFtools, variants with genotype quality less than 20 and with more than 50% missing data were excluded, then individuals with more than 50% missing data were removed. The resulting set of variants were filtered to retain only biallelic SNPs. Then removed variants which showed more than 80% of observed heterozygosity (Pearman et al. 2022). This filtering approach removes loci that are strongly deviating from HWE. Deviating from HWE heterozygosity expectations can arise from factors such as repetitive genomic elements and paralogs. Following this filtering we obtained the *globally non-thinned dataset* with 226,197 SNPs from 80 individuals.

Environmental data

For each sampling location 19 bioclimatic variables (Booth 2022) were downloaded from WorldClim database (Fick & Hijmans, 2017) at 2.5-minute resolution using the latitude and longitude coordinates using the R\raster package in R (Hijmans & van Etten, 2012). These variables are listed along with their abbreviations in Table 3 - S2.

Population genetic structure

Admixture (Alexander et al., 2009) was used to infer the population genetic structure using the SNP set referred to as the *globally thinned dataset*. Admixture was run with a with a major termination criterion of 1×10^{-9}, 1,000 bootstraps and 10-fold cross-validation for $K = 2–10$, where K equals the number of genetic groups. The K that produced the lowest cross-validation error was selected as the best K value. Admixture results were visualised using R\Pophelper 2.3.1 package (Francis, 2017) and pie charts. To further complement the population structure analysis a principal component analysis (PCA) was done using the R\SNPRelate package (Zheng et al., 2012). R\Poppr package (Kamvar et al., 2014) was used to detect if any clonal genotypes were present in the samples.

Genetic diversity

For each population we calculated the standard measures of genetic diversity - expected heterozygosity (H_E), observed heterozygosity (H_o), inbreeding coefficient (F_{IS}) and allelic richness (A_R) as average values across loci using the *globally thinned dataset*. The 95% confidence intervals of F_{IS} and A_R were calculated with 1000 bootstraps. These were calculated using the R\diveRsity package (Keenan et al., 2013).

Selfing rates of populations

Using the SNP set in the *globally thinned dataset,* selfing rates of populations were calculated using the R\inbreedR package (Stoffel et al., 2016). A measure of identity disequilibrium (the difference between joint identity by descent and the product of the separate probabilities of identity by descent for two loci, g2) was calculated for each population using the g2_snp function. Then population selfing rates were calculated using the equation presented in (David et al., 2007), where the population selfing rate = (1 + 5 * g2 - sqrt(1 + 10 * g2 + 9 * g2 ^ 2)) / (2 *g2).

Isolation by distance and environment

SNPs from the *globally thinned dataset* were used to calculate pairwise F_{st} across all populations using the R\DartR package (Gruber et al., 2018). Pairwise geographic distance between populations was calculated using the R\geosphere package (Hijmans et al., 2017) from longitude and latitude coordinates. Pairwise correlations between the 19 bioclimatic variables were calculated and one of each pair of highly correlated variables (Pearson's correlation coefficient > 0.70) were removed (Figure 3 - S1). This reduced the bioclimatic variables from 19 to five (mean diurnal range, bio2; isothermality, bio3: max temperature of the warmest month, bio5; mean temperature of the driest quarter, bio9; and annual precipitation, bio12). These five environment variables were first scaled and centred to account for the differences in magnitude. Then a measure of environmental distance between each population pair was calculated by pairwise Euclidean differences between the sites. Mantel tests were used to test the associations between linearized F_{st} calculated as (F_{st} / (1- F_{st}) with geographic distance and environmental distance using the R\ade4 package based on 999 permutations (Dray & Dufour, 2007). A partial

Mantel test was used to test the associations between linearized F_{st} calculated as (F_{st} / (1- F_{st}) with environmental distance after controlling for the geographic distance using R\ecodist (Goslee & Urban, 2007) package based on 999 permutations.

Genetic environment association

The genetic environment association (GEA) analysis was performed using BayPass V2.2 (Gautier, 2015). BayPass can identify loci under selection using an F_{ST}-like statistic (XtX) (Günther & Coop, 2013) that explicitly accounts for the confounding effect of population structure. To do so, BayPass first estimates the covariance matrix of population-level allele frequencies (Ω matrix in Gautier, 2015), which can capture even complex demographic histories. Then BayPass detects outlier loci among variants exhibiting the highest XtX values. We also used BayPass to identify variants associated with population-specific covariates (climate variables) (Gautier et al., 2018; Leroy et al., 2020). The program estimates Bayesian factors (BF) by comparing models with and without covariates. Using SNPs found in the *globally non-thinned dataset*, SNPs found on the first 20 longest scaffolds were imputed using BEAGLE (Browning & Browning, 2007) and the imputed data consisted of 223,912 SNPs. From this imputed SNP set, 5000 SNPs located outside annotated genes, were randomly sampled, and pruned for linkage disequilibrium using a window size of 50 kb, a step size of 5 bp and an r^2 threshold of 0.5 using plink V1.9 (Purcell et al., 2007). First, we ran BayPass using default parameters for the core model to estimate the covariance matrix of allele frequencies (Ω matrix) using this random SNP set. To identify allele frequency variation associated with environmental variables BayPass was run again by activating the covariate mode by providing the covariate file (-efile flag) and using the total SNP set obtained after imputing the *globally non-thinned dataset*. The covariate file consisted of the latitude, longitude and the 19 bioclimatic variables for each of the eight populations. The Ω matrix was provided with the -omegafile flag and -scalecov flag was provided to scale the covariates. Then genome-wide XtX and Bayesian factors were analysed in non-overlapping 30kb windows using the windowed-Z analysis (WZA) (Booker et al., 2021), with the top 5% of windows designated outliers in R. We termed windows that were in the top 5% of XtX and the top 5% of BF as climate adaptation candidate windows. In order to decide on a window size, first the top 1% XtX outlier SNPs were identified, then genome wide XtX were analysed in different sizes of non-overlapping windows (20kb, 30kb,

50kb) using WZA, and the window size that contained the greatest proportion of SNPs from the top 1% XtX in the top 5% WZA windows was selected.

We conducted a gene enrichment analysis using R\topGo package (Alexa & Rahnenfuhrer, 2006) after extracting the gene ontology terms associated with SNPs found in the XtX outlier windows that also overlapped with outlier windows for any one of the 19 bioclimatic variables (i.e., climate adaptation candidate windows).

Gradient forest (GF) predictions

To select candidate SNPs to be used in this analysis the top 1% XtX outlier SNPs that were also in climate adaptation candidate windows were identified. From them one SNP per window with the most significant association was selected. A gradient forest analysis using the R\gradientforest package (Ellis et al., 2012) was used to determine which environmental variables best explained the genetic variation. Gradient forest is a regression tree-based machine learning algorithm that searches for breaks across continuous environmental predictors. The model was run with 500 trees and a correlation threshold of 0.5. To visualise the gradient forest predictions across the range of the populations where the individual samples were collected, bioclimatic variables for 100,000 randomly generated geographic coordinates were extracted. These climatic variables were transformed for each random point using the gradient forest model above. These values were summarised using a principal component analysis (PCA). As suggested in the gradient forest manual, the top three principal components were transformed for visualisation, where the red value is defined by PC1+PC2, the green value by –PC2, and the blue by PC3+PC2-PC1.

Genetic offset calculation

The gradient forest analysis was extended to predict “genetic offset” following the methods presented in (Fitzpatrick & Keller, 2015). As the genomic offset analysis does not control for population structure, we used the climate adaptation candidates as input, as has been suggested in the literature (Láruson et al. 2022).Here genetic offset is a measure of how much locally adapted alleles are predicted to be perturbed from their present frequencies within a population for a given amount of environmental change. This can forecast the geographic regions with high

genetic mismatch under future climates. First the gradient forest model was recalculated using a subset of six uncorrelated bioclimatic variables (bio15, bio3, bio4, bio13, bio9 and bio5). These were selected by moving down the list of ranked importance for the full model above and discarding variables highly correlated (Pearson's r>0.7) with a variable of higher importance. Using 100,000 random points generated across the sampling range current and projected future values for these variables were extracted. Projected future values were extracted for the year 2050, under two RCP (Representative Concentration Pathway) greenhouse gas emission pathways: RCP45 and RCP85. For each random point both current and predicted bioclimatic variables were transformed based on the importance in predicting the genomic variation using the gradient forest model. Then the genetic offset was calculated as the Euclidean distance between these current and projected future values. Distribution of genetic offset along the geographic locations were plotted using the R\viridis package (Garnier et al., 2021).

Climate adapted loci and gene retention following whole genome duplication.

First, we calculated the number of ancestral (pre duplication) genes as those orthologous genes present in *Sorghum bicolor* and in one or both of the two subgenomes of *B. decipiens* using two independent scanning methods MCScanX (Wang et al., 2012) and OrthoFinder (Emms & Kelly, 2019). The division of *B. decipiens* reference genome into sub genome A and sub genome B was done as a part of chapter two of this thesis using differences in repeat abundance that were reflected in differing kmer distributions. Then a list of genes found in the climate adaptation candidate windows were extracted. Then we counted the number of genes that were present as single copy and were adaptation candidates, genes present as duplicates and one or both are adaptation candidates, genes present as singles copy and not adaptation candidates and genes present as duplicate and were not adaptation candidates. We then used a two-sided Fisher's exact test to test the null hypothesis of no association between copy number (single/duplicated) and adaptation status.

Results

Population genetic structure

In the Admixture run the K that produced the lowest cross-validation error was k=8 (Figure 3 - 2(A), Figure 3 - S2). Genetic clustering of individuals largely reflected their geographical distributions. Three of the four NSW populations shared the same Admixture clusters (mainly turquoise, and light green). However, the most southern population (CamdenN) was almost exclusively composed of the dark blue cluster. Similar geographic structuring was also identified in Queensland with the medium greens and yellow and light blues clusters found in this region. Interestingly, although some populations were composed of multiple clusters, little evidence of admixture was identified (e.g., BarabbaN, MuswelN), while in other cases evidence of admixture between clusters was apparent (e.g., HazeldeanQ, LakecalQ). The HazeldeanQ population also appeared to contain clusters found in both NSW and Queensland (Figure 3 - 2(A)(B)).This pattern of genetic clustering was also observed in the PCA analysis (Figure 3 - 2(C)), where individuals from NSW and Queensland populations formed separate clusters, but individuals from the HazeldeanQ population were found in the middle of the PCA, intermediate between these groups, while some individuals in this population are clustered with populations from NSW (Figure 3 - 2(C)). No clonal genotypes were detected among the sampled individuals.

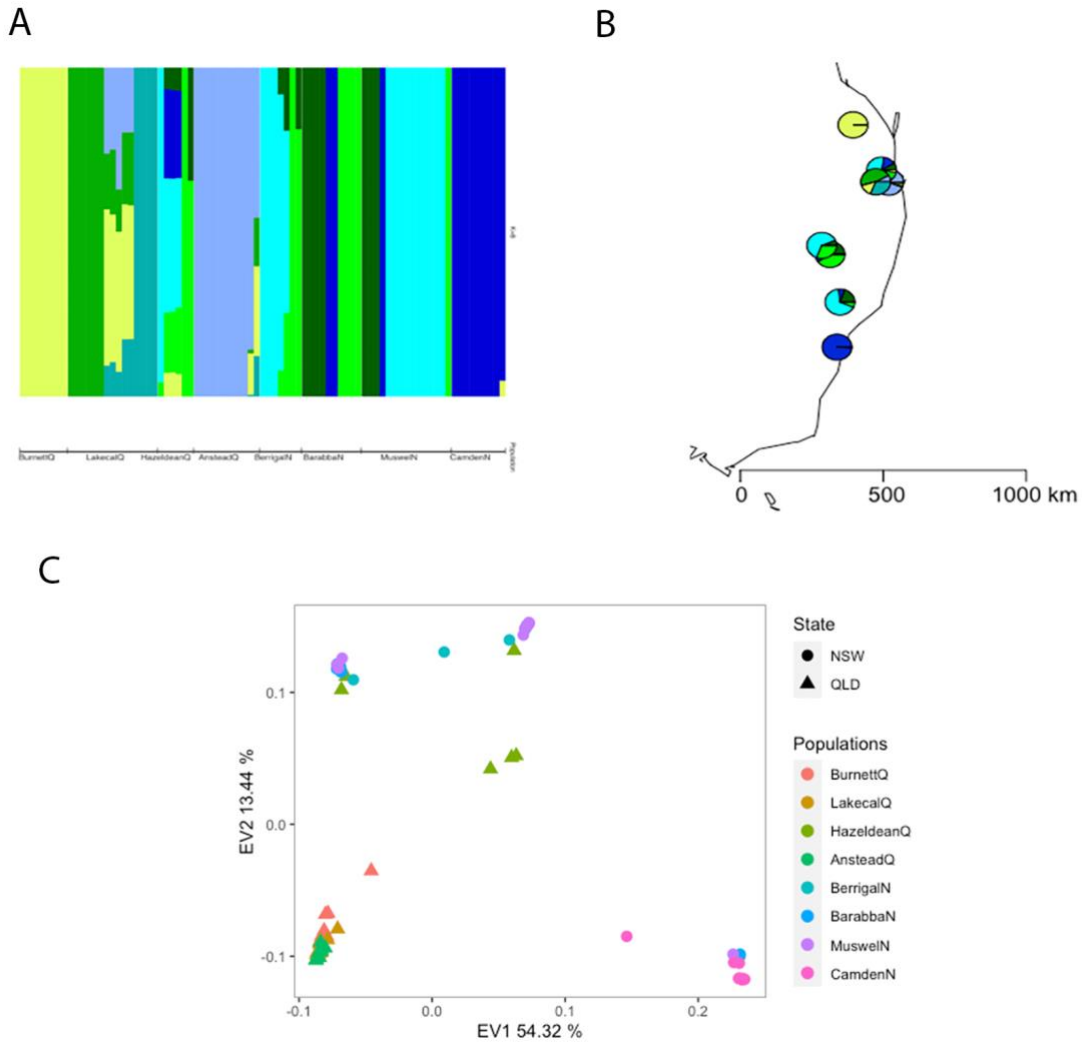


Figure 3 - 2. Genetic differentiation of sampled populations of *B. decipiens*. (A) A distruct plot showing the results of Admixture run of the globally thinned dataset for k=8. Populations from North to South are arranged left to right in the plot. (B) Population pie charts for k=8 mapped to represent their geographical distribution, admixture proportions for each population are displayed, colours correspond to the clusters in the distruct plot. (C) Principal component analysis of the globally thinned dataset, first two eigenvectors are presented, individuals are coloured according to their population.

Genetic diversity and Selfing rates of populations

Average values of calculated standard measures of genetic diversity were as follows- H_E - 0.18 (range 0.03-0.32), H_o - 0.18 (range 0.01-0.42), F_{IS} - 0.01 (range -0.20-0.52) and A_R - 1.24 (range

1.07 -1.72). The calculated population selfing rates using inbreedR for the eight populations ranged from significant self-fertilisation to complete outcrossing with an average estimate of 0.36 (range 0.004-0.79). Calculated genetic diversity statistics and selfing rate estimates for each population is presented in Table 3 - 1.There is a substantial variation 0f the distribution of these diversity measures among the populations. Some populations with larger values for expected heterozygosity also have negative values for inbreeding coefficient (BerrigalN, CamdenN, HazeldeanQ, MuswelN). BarbbaN had a high expected heterozygosity along with a high inbreeding coefficient. This population also had a higher estimated selfing rate.

Table 3 - 1. Summary of the calculated standard measures of genetic diversity per population averaged across loci and estimated population selfing rates.

Population	**H_E**	**H_o**	**F_{IS}**	**F_{IS} 95% CI**	**A_R**	**A_R 95% CI**	**Estimated selfing rate**
AnsteadQ	0.03	0.01	0.06	[0.05, 0.06]	1.07	[1.00, 1.11]	0.23
BarabbaN	0.22	0.05	0.52	[0.44, 0.56]	1.54	[1.01, 1.74]	0.79
BerrigalN	0.30	0.42	-0.20	[-0.24, 0.06]	1.67	[1.55, 1.69]	0.55
BurnettQ	0.05	0.02	0.07	[0.04, 0.08]	1.11	[1.02, 1.13]	0.004
CamdenN	0.17	0.17	-0.15	[-0.16, -0.12]	1.17	[1.06, 1.23]	0.01
HazeldeanQ	0.28	0.35	-0.17	[-0.51, 0.02]	1.65	[1.46, 1.70]	0.39

LakecalQ	0.05	0.01	0.10	[0.09, 0.10]	1.10	[1.01, 1.15]	0.59
MuswelN	0.32	0.42	-0.18	[-0.43, 0.05]	1.72	[1.62, 1.82]	0.34

Isolation by distance and environment

Linearized pairwise genetic distance between locations (F_{st} / (1- F_{st}) was significantly correlated with geographic distance, suggesting signals of isolation by distance (Mantel's r = 0.58, *P*-value=0.01). Linearized pairwise genetic distance between locations (F_{st} / (1- F_{st}) was not correlated with environmental distance, suggesting absence of isolation by environment (Mantel's r = 0.21, *P*-value=0.12). There was no significant association between the genetic distance and environmental distance after controlling for the geographic distance using the partial mantel test (Mantel's r=0.22, *P*-value = 0.18), suggesting the absence of isolation by environment among the sampled populations.

Genetic environment association

There were 720 XtX outlier SNPs (top 1%) in 692 climate adaptation candidate windows (top 5% WZA windows for XtX and BF). GO terms related to abiotic and biotic stress responses such as defence response, cellular response to water deprivation and salt stress were among the top five most significant terms associated with 849 genes identified in climate adaptation candidate windows (Table 3 – 1), further supporting the role of these genomic regions in local adaptation.

Table 3 - 2. Summary of the first five most significantly overrepresented GO terms among genes present in climate adaptation candidate windows relative to background genes.

GO ID	Term	Annotated	Significant	Expected	P-value
GO:0006952	Defence response	4570	70	42.64	1.2×10^{-6}
GO:0042631	Cellular response to water deprivation	247	12	2.30	3.4×10^{-6}
GO:0071472	cellular response to salt stress	217	10	2.02	3.5×10^{-5}
GO:0043447	alkane biosynthetic process	9	3	0.08	6.1×10^{-5}
GO:0097167	circadian regulation of translation	57	5	0.53	1.8×10^{-4}

Gradient forest predictions

We used 345 candidate SNPs for the gradient forest prediction analysis which represented one candidate SNP per candidate climate adaptation window (692 windows). The SNP with the highest association was selected according to the p-value. Based on the gradient forest analysis bioclimatic variables representing temperature and precipitation, particularly minimum temperature of warmest month (bio6), precipitation of the warmest quarter (bio18), precipitation seasonality (bio15), mean temperature of the coldest quarter (bio11) and precipitation seasonality (bio10) were the top five important variables that most strongly explained the genetic variation (Figure 3 - 3(A)). Differences in environmentally associated genetic variation

can be seen across the range (Figure 3 - 3(B)). This difference was apparent between the populations from the two states, NSW, and Queensland.

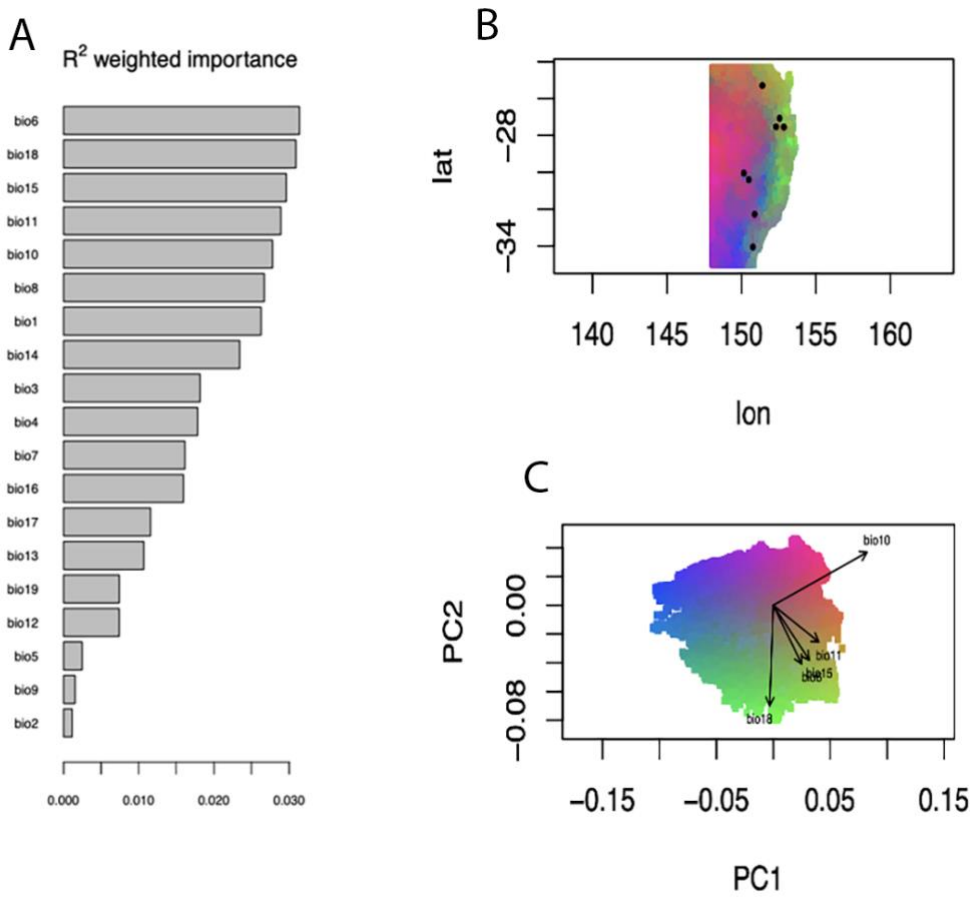


Figure 3 - 3. Summary of the gradient forest predictions. (A) Ranked importance of environmental variables on the genetic composition of the populations based on the gradient forest analysis. (B) Gradient forest transformed climate variables plotted on a map using the randomly generated geographic coordinates. Similar colours reflect similar allele frequencies at climate adaptation loci. Colours are based on principal components analysis (C) of transformed climate variables. Points on map reflect sampled locations. (C) Plot of the principal component analysis of the transformed climate variables representing the colours of the map (B).

Genetic offset calculation

A higher genetic offset was seen towards more inland locations while genetic offset decreased when moving towards the northern coast. As expected, genomic vulnerability increased as the climate change scenarios became more severe (Figure 3 - 4(A) (B)).

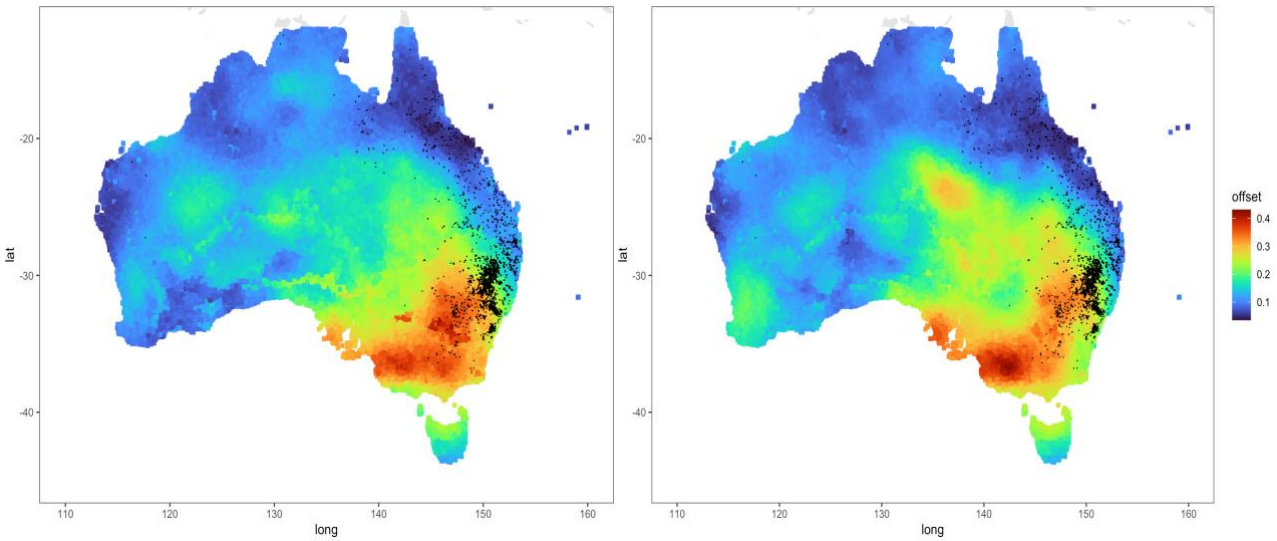


Figure 3 - 4. Genetic offset across geographic locations for future climate change scenarios. (A) Genetic offset based on 2050 RCP45 projections. (B) Genetic offset based on 2050 RCP85 projections. Species distribution is seen as black dots on each map representing occurrence records from Atlas of living Australia (Belbin et al. 2021).

Climate adapted loci and gene retention following whole genome duplication.

There were 849 genes found in the 692 climate adaptation candidate windows (related to 720 XtX outlier SNPs). We found the proportion of genes that were identified as climate adaptation candidates and duplicated in the genome is higher than the proportion of genes that are climate adaptation candidates and single copy in the genome (McScanX, Fisher's exact test, *P*-value = 1.52×10^{-10} and OrthoFinder, Fisher's exact test, *P*-value = 2.67×10^{-5} (Table 3 – 3). We also assessed if there was an excess of climate adaptation candidates in one subgenome over the other, but did not find a significant association (McScanX, Fisher's exact test, *P*-value = 0.61 and OrthoFinder, Fisher's exact test, *P*-value = 0.38) (Table 3 - S5).

Table 3 - 3. Summary of the number of adaptation candidate genes retained as duplicated or single copy following the most recent allopolyploidy event identified in *B. decipiens*.
The ancestral gene number is the number of orthologous genes shared between *sorghum* and *B. decipiens* and were in sequenced windows. Duplicated genes are retained in both subgenomes. The genes that reverted to single copy are in the A or B subgenome only. Proportion of single copy adaptation candidates out of total single copy genes and proportion of duplicated adaptation candidates out of total number of duplicated genes is given in parenthesis.

Clustering Method	**Ancestral gene number (sequenced)**	**Total Single copy genes+adaptation candidates**	**Total Genes retained as duplicated and one or both are adaptation candidates**	**Total Single copy genes + non-adaptation candidates**	**Total Duplicated genes + non-adaptation candidates**
McScanX	19, 523	122 (0.02)	388 (0.03)	7,115	11,898
OrthoFinder	15, 262	76 (0.01)	222 (0.02)	5,563	9,401

Discussion

Using landscape genomics, we discovered evidence of climate adaptation in *Bothriochloa decipiens*, especially for loci related to abiotic and biotic stress response. This is one of the few genomic studies of an Australian understory species, and these findings suggest that climate adaptation should be a consideration when grassland restoration efforts involving this species are implemented. Genomic offset analysis identified populations in the southwestern portions of

the range as more vulnerable to maladaptation under climate change. Evidence of strong neutral population structure coupled with sometimes high rates of self-fertilisation suggests local pollen and seed dispersal. This suggests that populations will be reliant on trait plasticity or *in situ* evolutionary adaptation to local conditions rather than natural migration among populations to respond to changing climates. However, we discovered that a recent whole genome duplication event was likely a significant source of genetic diversity facilitating local adaptation to climate, as candidate climate adaptation genes were more likely to be retained as duplicated. These findings are relevant to the planning of restoration management efforts that consider the genetic diversity and evolutionary potential of this species and involve assisted gene flow to mitigate maladaptation under future climate change. However these findings should be further supported by increasing the sampling efforts as they are based on a limited number of populations.

Signals of climate adaptation

Landscape genomic approaches are becoming increasingly efficient in identifying adaptive divergence in non-model species (Fitzpatrick & Keller, 2015; Forester et al., 2016; Jordan et al., 2017; Rellstab et al., 2015). Using these approaches, we identified strong signals of climate adaptation in *B. decipiens*. Overall, 32% of 2239 XtX loci also were correlated with climate variables making these 720 loci strong candidates for targets of climate mediated selection. Also, these genomic regions contained an excess of genes related to biotic and abiotic stress response, including defence response, response to water and salt stress. Among top candidate genes with annotations related to drought stress response was a homolog of bHLH112 (basic helix loop helix) transcription factor. In plants some members of the bHLH family of transcription factors, such as bHLH112, have been shown to be involved in abiotic stress tolerance, including drought stress (Babitha et al., 2015; Liu et al., 2014; Li et al., 2021). We also identified a homolog of the Cys2/His2 zinc finger transcription factor (ZAT18), and it is known to be a positive regulator of plant drought stress tolerance (Yin et al., 2017). These findings further corroborate the putative role of these regions in local adaptation in this species. The exposure of *B. decipiens* to strong aridity gradients across the range (Figure 3 - 1) is a potential driver of these signals of climate adaptation, especially given the prominence of putative drought response genes in climate adaptation candidate windows. The importance of climate variation, and aridity in particular, in structuring adaptive genetic variation has also been identified in several other C4 grasses (Gould et al., 2018; Gray et al., 2014; Guo et al., 2021).

Local adaptation to climate via allopolyploidy

Bothriochloa decipiens has experienced a relatively recent whole genome duplication (WGD ~6 MYA) and is a diploidized allotetraploid species with stress response genes overrepresented among the duplicated genes in the genome (Chapter 2). We also discovered that genes retained as duplicated following the WGD event were more likely to be climate adaptation candidates than those reduced to single copy. As neofunctionalization and specialisation can be possible fates of genes retained as duplicates, retention of such genes may promote adaptive evolution to biotic and abiotic stresses (Cheng et al., 2018; Defoort et al., 2019; Jiao et al., 2018; Panchy et al., 2016; Ren et al., 2018; Soltis & Soltis, 2016). Our findings support the hypothesis that gene duplication caused by allopolyploidy in grasses has contributed to their ability to adapt to new environments, thereby enabling successful expansion and establishment in novel climates (Godfree et al., 2017; Linder & Barker, 2014;(Ahrens et al., 2020). In switchgrass, the subdominant genome was found to contribute more to SNP-heritability in common gardens (Lovell et al., 2021). They suggested this was due to the relaxation of purifying selection in this subgenome, which provided the underlying evolutionary novelty for adaptation. We looked for evidence of a bias in the contribution of the subgenomes to climate adaptation but failed to detect any evidence for this.

Predictions of future maladaptation

Heterogeneous patterns of future maladaptation for climate were seen throughout the species distribution. Gradient forest analysis indicated that bioclimatic variables representing both precipitation and temperature were the most important variables associated with genetic variation. Calculating genetic offset can predict the possible fate of populations across a species range when faced with predicted climate change (Borrell et al., 2020; Ingvarsson & Bernhardsson, 2020; Pina-Martins et al., 2019) and has been suggested to be used in conservation management (Capblancq et al., 2020). Our genetic offset analysis revealed that some populations *B. decipiens* may be more adversely impacted by climate change than others and if the genetic diversity of these populations is low, this means the potential of these species to adapt to future climate change will be limited (Willi et al. 2006; Hoffman et al. 2014). A pattern of increasing genetic offset is seen when moving towards more inland locations, particularly in the south of the range. Similar findings, where inland populations become more

vulnerable to climate change, were reported in the grass species *Themeda triandra* (Ahrens et al., 2020) but this was seen only when ploidy was excluded in the model. When ploidy was included the vulnerability of the inland populations reduced. To keep pace with climate change the high-risk regions with higher genetic offset would likely need to adapt more otherwise populations in these regions may decline.

Further there are some important issues to be considered in using genetic offset for predicting maladaptation to climate in a species. This method assumes that the relationship between environment and allele frequency is at a local adaptation equilibrium, and will remain constant beyond present climate and genetic background. There is no guarantee that these conditions will remain the same. As species will encounter environments different from the present conditions in the future, genetic offset predictions into these regions should be looked at with caution (Rellstab et al. 2021). Genetic offset predictions are based on SNP allele frequencies used to identify local adaptation and allele frequencies across a species range can vary due to demographic history unrelated to local adaptation (Hoffmann et al. 2021). As a result the predictions may be based on false positives climate candidates. Further the climate change predictions that the genetic offsets rely on are also uncertain. Differences in the genetic background among populations and the interactions of fitness affecting loci are unknown and may influence the accuracy of these predictions. Indeed, evolutionary fates of populations are influenced by many interrelated factors that can be difficult to predict (Reside et al., 2018). However, despite these uncertainties, developing means to integrate evolutionary forces in species' response to climate change predictions, such as the genetic offset analysis, may still assist in conservation and management of populations (Allendorf et al., 2010; Brown et al., 2016; Waldvogel et al., 2020).

Population genetic structure

The species is known to be frequently cleistogamous in nature, that is, its ovules are self-fertilised with pollen inside flowers that normally do not open (Connor, 1979). Long days, however, have been reported to cause culm elongation leading to chasmogamous flowers (Connor, 1979). Surprisingly, our findings support high selfing rates in only some populations (particularly at the margins of the sampled range) as mixed mating and even high levels of outcrossing were also observed. Together with a lack of substantial clonal reproduction, this

suggests that reproduction in this species involves wind pollination in many populations. Mixed mating wind pollinated species are known to be relatively rare in nature (Friedman & Barrett, 2009), and future studies examining the cause of this geographic variation in selfing rates (e.g., plant density, the frequency of chasmogamy/cleistogamy) should be undertaken. Mixed mating systems with intermediate and variable selfing rates are predicted to be more stable evolutionarily (Goodwillie et al., 2005). Further this mode of reproduction can provide resilience to climate change perturbations through compensatory reproductive responses (Jones et al., 2013), and in wind pollinated species self-fertilisation may be particularly important when local plant densities decline.

Overall, our analysis of population structure identified several differentiated genetic clusters. Selfing plants have lower genetic diversity within populations and higher differentiation among populations (Koelling et al., 2011; Song et al., 2006; Sweigart & Willis, 2003; Williams et al., 2001) due to reduction in pollen migration and high levels of inbreeding contributing to greater drift (Culley & Wolfe, 2001; Duminil et al., 2009). The presence of high selfing in some populations likely has contributed to the substantial population structure we observed (mean F_{st} = 0.42). We also identified substantial mixing among genetic clusters in some populations, and this included the presence of admixed individuals. This may suggest that assisted migration among differentiated populations may not pose significant fitness costs, although this should be investigated further by looking at the fitness effects of interpopulation crosses.

Populations also showed significant isolation by geographic distance. Isolation by distance is caused by local accumulation of genetic differences due to dispersal limitations caused by geographic restrictions and is common in plant populations (Meirmans, 2012; Slatkin, 1993). However, we did not see signals of isolation by environment in neutral genetic markers. Greater signals of IBD compared to IBE are often identified in plant studies (Sexton et al., 2014; Shafer & Wolf, 2013). Lack of IBE in neutral genetic markers does not necessarily mean lack of local adaptation to environment and examples of climate adaptation despite limited population structure abound in the literature (Andrew et al., 2012; Dionne et al., 2008; Sambatti & Rice, 2006). Local adaptation will frequently involve only a subset of regions in the genome while the environment can modulate gene flow and drift leading to IBE genome wide. It is possible that IBE might be present in this species but weak and given the small number of populations used in the study and the covariation between environmental and geographic distance, our power is limited. Future studies with a greater number of populations across a greater range of

environmental variation are warranted. However, our preliminary results suggest that a larger scale sequencing effort will be worthwhile.

Management recommendations

Grasslands are one of the largest terrestrial biomes and hotspots of biodiversity. They provide important ecosystem services such as carbon sequestration, pollinator promotion, erosion control, water supply and regulation and food and fodder (Bengtsson et al., 2019; Zhao et al., 2020) along with high conservation value (Dengler et al., 2014). Large scale destruction of grasslands has occurred due to land use change for agriculture and settlements, climate change and invasive species (Bakker & Berendse, 1999; Steffen et al., 2015; Walther et al., 2009). Consequently, taking management decisions that ensure climate change resilient restoration of degraded grasslands is important in an ecological and economical context. *Bothriochloa decipiens* is recognized as an important species for grassland rehabilitation due to its ability to establish well from direct seeding on many soil types, and the ability to withstand pressure caused by overgrazing (Simon & Alfonso, 2011). Our results provide the critical first step in assessing the suitability of using *B. decipiens* in grassland restoration by practising assisted gene flow among populations for climate change resilient restoration. The use of this species to practise assisted gene flow to facilitate climate adaptation can be justified by the presence of signals of strong climate adaptation especially in relation to drought stress. Further the strong population structuring and isolation by distance we identified suggests local dispersal and points to the importance of assisted gene flow as natural gene flow may limit the rate of adaptation to climate change in some instances. Simulations parameterized using population genomic data such as these, would shed light on the importance of natural versus assisted migration in maintaining populations in the face of climate change.

Genetic offset analysis revealed that populations towards more inland locations were more genetically vulnerable under future climate change. Mixing seeds from coastal populations with inland populations may be unlikely to help improve the resilience of these vulnerable populations, however, as these populations come from wetter regions it is unlikely, they contain the requisite adaptive genetic variation. Importantly, the sampling of populations across a wider range of aridity levels in future studies is needed as this may uncover important adaptive genetic variation that can be used to mitigate the impacts of future droughts in these regions.

Consequently, findings from this study should be further investigated by increasing the sampling effort.

Conclusion

We discovered evidence of climate adaptation in *Bothriochloa decipiens*, especially for loci related to abiotic and biotic stress response. Presence of strong neutral population structure and sometimes high rates of self-fertilisation suggests local pollen and seed dispersal. Genomic offset analysis identified populations in the southwestern portions of the range as more vulnerable to maladaptation under climate change. However, these predictions on maladaptation should be further investigated by increasing the sampling efforts, especially towards the more inland locations of the species distribution range. Then findings from this study provide the first step in assessing the suitability of *B. decipiens* to be used in restoration of Australian grasslands as species specific information on genetic diversity and signals of adaptation in grassland species is important in restoration.

Appendix II - Supplemental information for : Signals of climate adaptation and predictions of future maladaptation in an Australian native grass

Table 3 - S1. Summary of the abbreviations used to identify the populations used in this study from which individual samples were collected, the geographic location, state each population belongs to and the number of individual samples used per population.

Population abbreviation	**Latitude**	**Longitude**	**State**	**Number of individuals sampled**
AnsteadQ	-27.548	152.859147	Queensland	12
BarabbaN	-30.393709	150.485174	NSW	11
BerrigalN	-30.0392	150.1523	NSW	8
BuenettQ	-25.288891	151.4144	Queensland	8
CamdenN	-34.048427	150.767832	NSW	9
HazeldeanQ	-27.067943	152.569876	Queensland	7
LakecalQ	-27.53304	152.33292	Queensland	18
MuswelN	-32.271531	150.886235	NSW	15

Table 3 - S2. The list of the 19 bioclimatic variables downloaded from the WorldClim database used in the study along with their abbreviations.

Name of bioclimatic variable	**Abbreviation**
Annual mean temperature	bio1
Mean diurnal range	bio2
Isothermality	bio3
Temperature seasonality	bio4
Max temperature of warmest month	bio5
Min temperature of warmest month	bio6
Temperature annual range	bio7
Mean temperature of the wettest quarter	bio8
Mean temperature of the driest quarter	bio9
Mean temperature of the warmest quarter	bio10

Mean temperature of the coldest quarter	bio11
Annual precipitation	bio12
Precipitation of the wettest month	bio13
Precipitation of the driest month	bio14
Precipitation seasonality	bio15
Precipitation of the wettest quarter	bio16
Precipitation of the driest quarter	bio17
Precipitation of the warmest quarter	bio18
Precipitation of the coldest quarter	bio19

Table 3 - S3. Summary of the 19 bioclimatic variables for each of the populations used in the study and the range of each variable across the species range.

Population	bio1	bio2	bio3	bio4	bio5	bio6	bio7	bio8	bio9	bio10	bio11	bio12	bio13	bio14	bio15	bio16	bio17	bio18	bio19
AnsteadQ	20.02	12.04	52.06	410.98	30.29	7.16	23.13	24.65	15.40	24.65	4.50	1042	150	31	46.96	423	115	423	124
BarabbaN	15.74	14.48	50.06	551.79	30.12	1.18	28.93	22.33	16.33	22.33	8.66	778	115	42	32.11	276	149	276	154
BerrigalN	16.94	12.55	47.04	539.06	30.18	3.49	26.69	23.33	13.91	23.33	10.00	696	98	37	26.49	235	143	235	152
BuenettQ	19.89	12.62	51.83	438.04	30.76	6.41	24.34	24.75	13.96	24.75	13.96	737	114	27	47.88	306	93	306	93
CamdenQ	17.00	11.53	50.35	418.98	27.80	4.89	22.90	21.68	12.482	21.97	11.53	942	113	45	29.00	329	149	291	170
HazeldeanQ	19.66	10.48	48.57	405.88	29.64	8.07	21.57	24.19	15.02	24.19	14.18	987	133	32	43.82	395	122	395	127

Lakeca lQ	19.68	12.75	52.36	436.90	30.78	6.42	24.36	24.59	13.80	24.59	13.804	832	118	26	48.23	348	96	348	96
Muswe lN	17.58	13.26	49.92	499.70	30.92	4.35	26.56	23.59	11.17	23.59	11.17	662	92	33	29.86	228	112	228	112
Range	27.35-11.26	16.25-5.1	82.08-39.24	657.58-63.29	40.42-22.98	22.05-2.53	32.18-10.04	31.81-7.83	26.21-5.55	31.95-17.11	25.94-4.82	2439-248	592-27	71-0	129.39-9.12	1409-78	240-1	1115-78	332-1

Table 3 - S4. Population pairwise FST matrix

	AnsteadQ	BarabbaN	BerrigalN	BurnettQ	CamdenN	HazeldeanQ	LakecalQ	MuswelN
AnsteadQ	0	0.45818397	0.51526903	0.40641584	0.8983359	0.51646158	0.31898328	0.44728795
BarabbaN	0.45818397	0	0.05032874	0.35965547	0.66630728	0.01556035	0.45672072	0.10563541
BerrigalN	0.51526903	0.05032874	0	0.4245157	0.60330325	0.00647402	0.51519429	0.0150127
BurnettQ	0.40641584	0.35965547	0.4245157	0	0.87474531	0.41161204	0.28057841	0.38293132
CamdenN	0.8983359	0.66630728	0.60330325	0.87474531	0	0.61602606	0.89336447	0.49480139
HazeldeanQ	0.51646158	0.01556035	0.00647402	0.41161204	0.61602606	0	0.50992859	0.02488067
LakecalQ	0.31898328	0.45672072	0.51519429	0.28057841	0.89336447	0.50992859	0	0.45425937
MuswelN	0.44728795	0.10563541	0.0150127	0.38293132	0.49480139	0.02488067	0.45425937	0

Table 3 - S5. Summary of the number of adaptation candidate genes retained in the two subgenomes A and B as duplicated or single copy following the most recent allopolyploidy event identified in *B. decipiens*. The ancestral gene number is the number of orthologous genes shared between sorghum and *B. decipiens* and were in sequenced windows. Duplicated genes are retained in both subgenomes. The genes that reverted to single copy are in the A or B subgenome only.

Clustering Method	**Ancestral gene number(sequenced)**	**Single copy genes + adaptation candidates**		**Genes retained as duplicated and one or both are adaptation candidates**			**Single copy genes + non-adaptation candidates**		**Duplicated genes + non-adaptation candidates**		
		A	**B**	**A**	**B**	**Both**	**A**	**B**	**A**	**B**	**Both**
McScanX	19,523	83	39	209	175	4	4407	2708	4875	3786	3237
OrthoFinder	15,262	39	37	127	93	2	3057	2506	3876	3540	1985

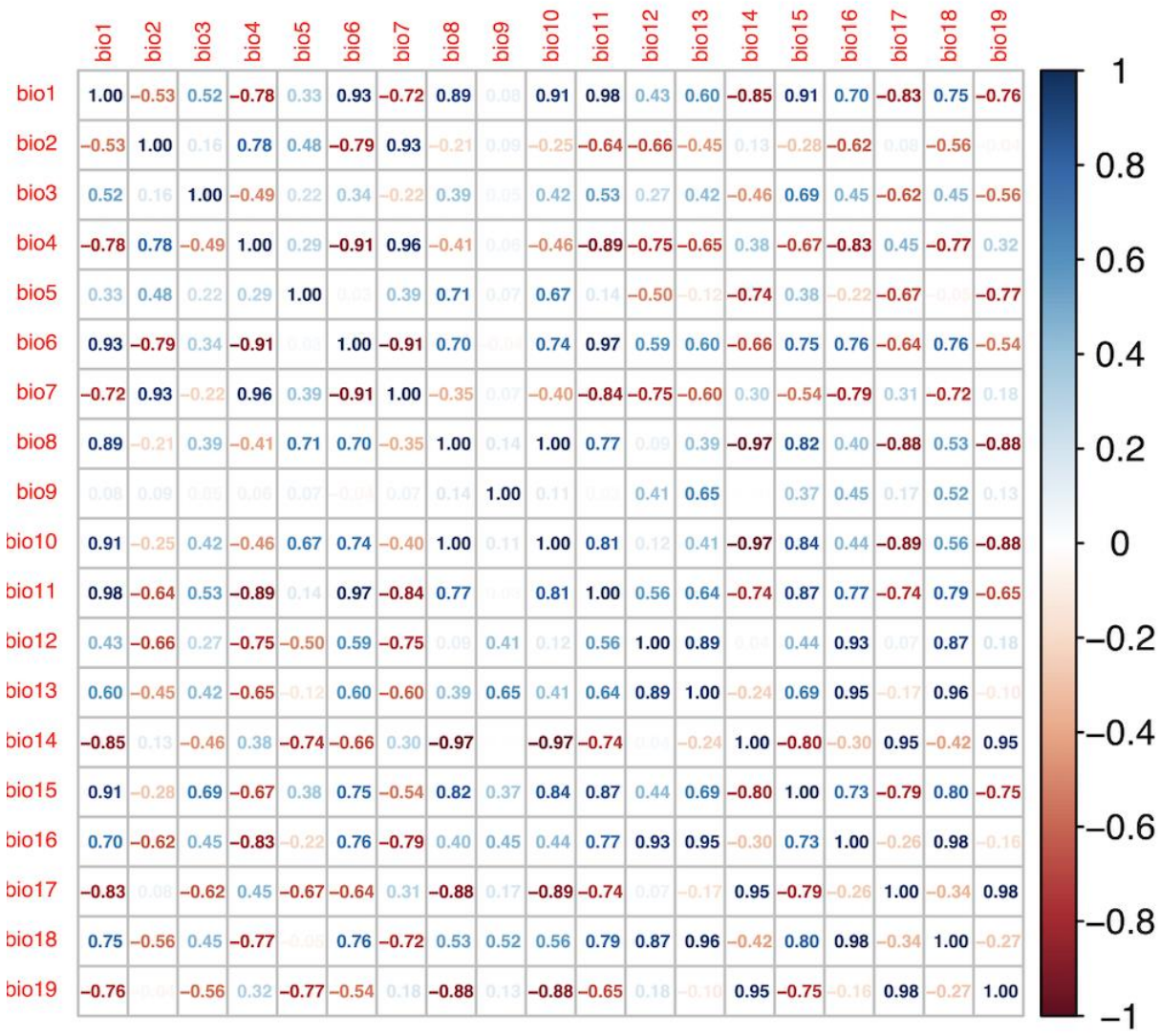
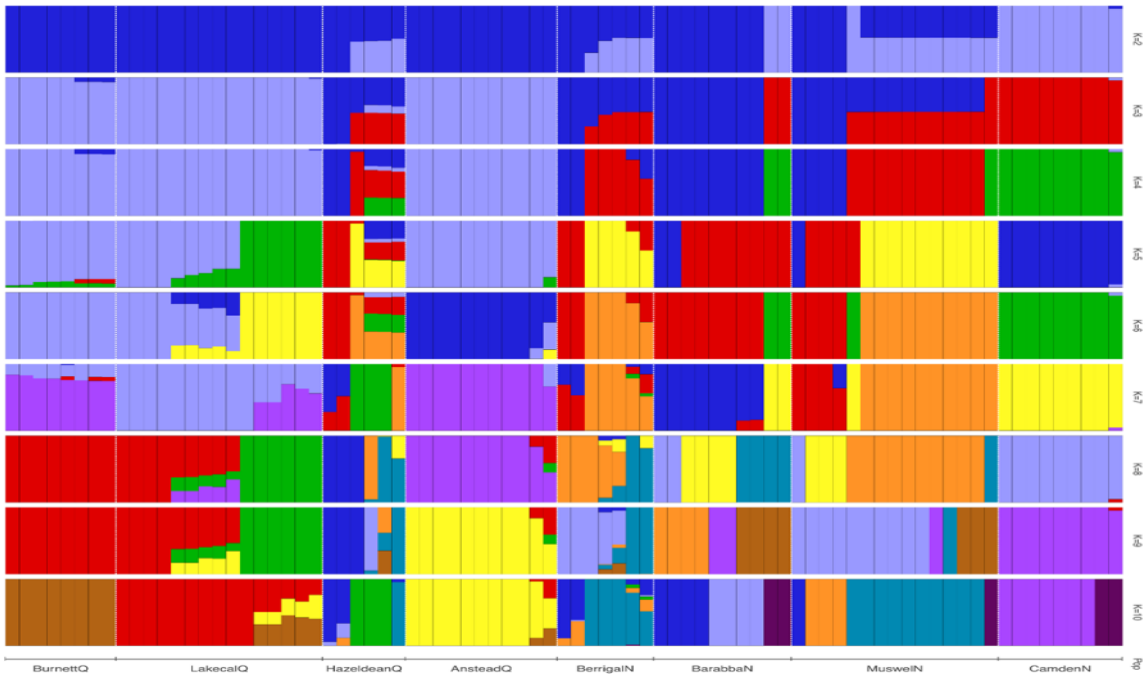
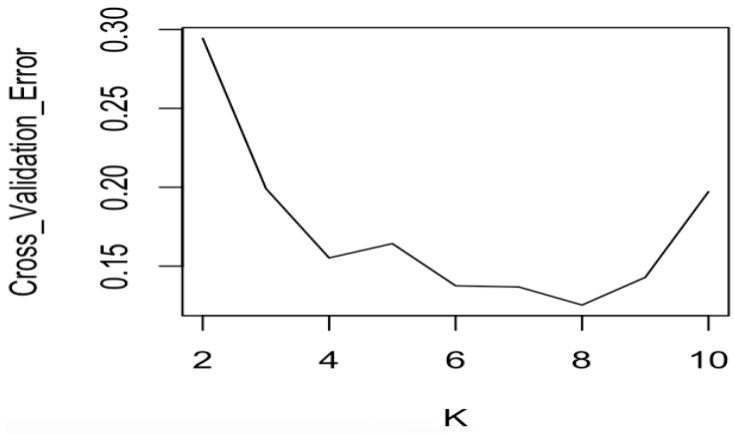


Figure 3 - S1. Corrplot showing pairwise correlation between the 19 bioclimatic variables.

A





B

Figure 3 - S2. (A) A distruct plot showing the results of Admixture run of the *globally thinned dataset* for k=2-10. (B) Plot of the distribution of the cross-validation score for each k value. Populations from North to South are arranged left to right in the plot.

Chapter 4 - Australian native grass *Bothriochloa macra* shows signals of local adaptation to drought.

Nissanka P. De Silva[1*], Christopher Lee[1], A. Fournier-Level[2], Joslin L. Moore[3] , Kathryn A. Hodgins[1]

[1]School of Biological Sciences, Monash University, Clayton, Victoria, Australia, 3800.

[2]School of BioSciences, University of Melbourne, Melbourne, Victoria, Australia, 3010.

[3]Arthur Rylah Institute for Environment Research, Heidelberg, Victoria, Australia, 3084.

[*]Corresponding author

In preparation for *Ecology and Evolution*

Abstract

Many grassland ecosystems have been heavily degraded by land use changes and are experiencing increased threats due to climate change. At present using ecological genomic approaches to improve climate change resilient restoration efforts of species is quite common, but studies that investigate the genetic diversity and local adaptation of grasses, especially in Australia, are limited. We explored patterns of genetic differentiation of traits and genome wide SNPs in an Australian native grass species *Bothriochloa macra* in relation to historic climates. We did this to assess if there was evidence consistent with local adaptation to climate in this key wide-ranging foundation species in the grasslands of south-eastern Australia. We used genotype by sequencing and identified 5,566 SNPs across 162 individuals (25 populations). Evidence of isolation by distance and by the environment was detected, indicating restricted dispersal that may also be influenced by the environment. We conducted a common garden experiment and measured early life history traits using seeds collected from 44 populations of this species. Germination and seedling characteristics were investigated because they represent the initial and most critical stage in establishment of a new generation of plants, hence play an important role in recruitment and population persistence. Among the measured traits, total biomass showed significant associations with climate variables related to drought, potentially indicating climate adaptation. Seedlings from high rainfall and low temperature areas yielded higher total biomass values when compared with seedlings from low rainfall and high temperature areas. This is consistent with the hypothesis that populations from more arid regions have evolved a 'drought tolerance' mechanism to survive drought by decreasing resource investment in overall growth and total biomass allocation. However, Q_{st}-F_{st} estimates of other seedling size characteristics were inconsistent with divergent selection. Preliminary findings from this investigation shed light into the suitability of using *B .macra* in climate change resilient grassland restoration and could help in planning potential management strategies involving AGF for restoration of degraded Australian grasslands.

Keywords – climate change, ecological restoration, grasslands, local adaptation, drought tolerance

Introduction

Ecological restoration of degraded ecosystems is a worldwide practice. Nearly $US 2 trillion per annum is invested in ecological restoration globally (Cunningham, 2008) and considering the rapid rate of habitat destruction worldwide, investment in ecological restoration will continue to increase further. Given its critical ecological importance and the large investment of time and money, a key concern is that it delivers long-lasting results particularly given that these efforts are carried out in an era of accelerated climate change. According to the report on climate change produced by IPCC in 2013 there will likely be a further rise of 2° in temperature above the present level of 1° C even if the CO_2 emissions are stopped completely by 2050. These pervasive environmental impacts suggest that existing strategies for ecological restoration should be assessed for their resilience to climate change. Consequently, when undertaking ecological restoration of degraded ecosystems, it may be prudent to select seeds/germplasm that will have a better chance of getting established and reaching reproductive maturity under climate change scenarios thereby contributing to revival of a healthy ecosystem (Prober et al., 2015).

Global climate change is already having far reaching effects on biodiversity (Hoffmann et al., 2019; Hoffmann & Sgrò, 2011; Scheffers et al., 2016) and such impacts are predicted to worsen (Nunez et al., 2019; Urban, 2015). Species can survive climate change either by migration to suitable climates or by *in situ* adaptation to new climates (Hoffmann & Sgrò, 2011). *In situ* adaptive responses can occur through adaptive phenotypic plasticity (Nicotra et al., 2010) or by natural selection of suitable genetic variants in the local population, i.e., genetic adaptation (henceforth referred to as adaptation). For such adaptive responses to occur the presence of genetic variability in ecologically relevant traits is essential (Sgrò et al., 2011). This can be aided through assisted movement of genotypes from other populations experiencing different climatic conditions across the landscape. However, before the assisted movement of genotypes is carried out to help populations adapt to changing climates, an understanding of both the population genetic structure as well as the degree of local adaptation to climate is key.

The transition from one life cycle stage to another is key to the long-term persistence and growth of plant populations (Fitter & Silvertown, 1983). The most important plant life cycle stage transitions are those associated with seed germination and seedling growth, as it is critical for the establishment of the next generation and also helps in colonization of new and existing

habitat (Schemske et al., 1994). Moreover, selection pressures are often highest during this regenerative phase of the life cycle (Harper, 1977). The specific environmental conditions in which the regeneration process occurs have been called the 'regeneration niche' of a species (Grubb, 1977). If the environmental conditions defining the regeneration niche move beyond the existing conditions, the establishment of the next generation will be affected and eventually this could cause local extinction unless the population rapidly adapts to the new conditions (Hudson et al., 2015). Plants are expected to be more sensitive to climate change during their early developmental stages than adult stages, representing a major bottleneck during establishment (Dalgleish et al., 2010; Fay & Schultz, 2009; Lloret et al., 2004).

Given the potential sensitivity of early life stages to their local environments, environmental heterogeneity may give rise to seedlings locally adapted to their climate of origin. A signature of such local adaptation is genetic differentiation of seedling traits among populations when measured in common gardens, especially if such trait variation is correlated with historic climates. Many studies have examined relationships between germination and seedling growth along the latitudinal, longitudinal, and altitudinal gradients of source populations (Ghildiyal, 2009; Loha et al., 2008; Mamo et al., 2006; Wu et al., 2019; Xiao et al., 2012). These geographic clines often reflect gradients in climatic factors such as precipitation and temperature and can reflect climate mediated divergent selection. Although patterns of divergence in seed and seedling characters observed can be due to genetic differences among individuals, maternal effects can also play a substantial role in governing variation for these traits in field collected seeds because of maternal seed provisioning (Aarssen & Burton, 1990; Lacey, 1998; Tungate et al., 2006; Wulff & Bazzaz, 1992).

The comparison between quantitative genetic differentiation (Q_{st}; Spitze, 1993) and neutral marker differentiation (F_{st}; Wright, 1951) can also be used indirectly to investigate if local adaptation might be involved in phenotypic divergence of populations (McKay & Latta, 2002; Merilä & Crnokrak, 2001; Spitze, 1993). Traits involved in local adaptation can be identified by comparison of population genetic differentiation estimated from putatively neutral genetic markers with population quantitative genetic differentiation estimated from quantitative phenotypic traits (Whitlock, 2008). If $Q_{st} > F_{st}$ for quantitative trait genetic differentiation among populations it is more likely due to divergent selection, $Q_{st} < F_{st}$ indicates stabilizing selection while $Q_{st} = F_{st}$ indicates neutral evolutionary divergence (Whitlock & Guillaume, 2009). Experimental and theoretical studies that compare molecular and quantitative genetic variation

reveal how natural selection has played a major role in shaping intraspecific variation in quantitative traits (Leinonen et al., 2008; Leinonen et al., 2013; McKay & Latta, 2002). However, well powered Q_{st}- F_{st} comparisons require substantial logistical effort as both quantitative genetic estimates of trait differentiation and estimates of molecular divergence for many loci and populations must be conducted, limiting the application of this approach (Whitlock 2008).

In addition to investigating the extent of climate adaptation, understanding population structure can be informative when assessing seed stocks that are used for restoration. Although assisted gene flow in restoration is mainly aimed at speeding climate adaptation (Aitken & Whitlock, 2013), epistatic incompatibilities (Schiffers et al., 2013) and maladaptation to non-climatic factors such as soil type (Wright et al., 2006) can cause reduced fitness after restoration efforts and substantial population differentiation may indicated the presence of such reproductive barriers among populations. For instance, ploidy can play an important role in shaping the differentiation among the population due to barriers against breeding among different cytotypes (Köhler et al., 2010). When human mediated transfer of propagules to assist dispersal and increase connectivity among populations by increasing the gene flow occurs, there is always the risk of maladaptation of the new genotypes and also the new genotypes might interbreed with the local population and produce offspring with lower fitness (outbreeding depression) (Hufford & Mazer, 2003). Therefore investigating the population structure of species used for ecological restoration is important to inform restoration strategies and to ensure that potential seed transfer occurs between populations with similar genetic structure and variability (Broadhurst et al., 2008). By contrast, significant population structure may also suggest that contemporary gene flow from pre-adapted populations may be insufficient to rescue populations from local extinction in response to rapid environmental change (DeSilva & Dodd, 2020), thereby further supporting human intervention.

One ecosystem under threat by both land use change in the past (Gibbons & Boak, 2002; Lunt & Morgan, 2002; Prober et al., 2015; Prober & Thiele, 2005) and pervasive climate change into the future (Pitman et al., 2007) are Australian grasslands. The Australian native grass species, *Bothriochloa macra* is an excellent candidate for restoration of Australian grasslands for several reasons. The species is relatively common and widespread in Australian grasslands. This distribution across four climatically diverse states in Australia means that local adaptation to

climate is likely. It is also relatively resilient to a number of biotic and abiotic stressors and is long lived (5-25 years). As an increaser species it benefits from grazing and modification of habitat by increasing in abundance making it widespread in overgrazed pastures (Whalley, 1977). It is highly drought tolerant (Semple et al., 1997) and its tolerance to weeds during establishment is greater compared to most other Australian native grasses (Hagon, 1977). The spread of this species along road verges is an indicator of the ability to establish well in degraded soils. Moreover, it has the ability to improve good soil properties by accumulation of organic matter (Moore, 1957). All these characters make *B. macra* a well-suited species for the restoration of degraded Australian grasslands. However, little is known about its population genetic structure and its degree of climate adaptation.

The main aim of this study is to determine whether there is evidence consistent with local adaptation to climate in the study species *B. macra* in relation to early life history traits such as seed germination and seedling growth. To do this, we examined if these traits showed any signatures of adaptation using common garden experiments and accessions sampled from across the species' range. We assessed its population structure using genome wide markers to determine the suitability of mixing genotypes derived from different sources and to assess evidence of divergent selection using a Q_{st}-F_{st} analysis. Specifically, we address the following questions: 1) Is there genetic differentiation of important early life history traits among populations and is this differentiation consistent with divergent selection? 2) Are these population differences associated with climate? 3) What are the patterns of population genetic divergence and genetic diversity within populations? This knowledge base will be key in developing future recommendations regarding the provenance of seeds to restore degraded grasslands and mitigating the potential negative impacts of climate change on establishment success.

Methods

Study species

Bothriochloa macra is a perennial tussock grass producing slender reddish flowering stems in summer and early autumn giving it the common name red grass or red-leg grass. *Bothriochloa macra* is a C4 species and active in warmer conditions with flowers produced from summer to

early autumn (Mitchell, 2002). The species is pollinated by wind and seed dispersal is either by wind or transportation through adhesion. According to the Atlas of Living Australia, *Bothriochloa macra* is distributed in QLD, NSW, VIC, and SA.

Sample collection

Seeds were collected from a total of 44 populations of *B. macra* (Figure 4 - 1). Field collections were done during January to March of the year 2019. Collections were timed to ensure the collected seeds were ripe. A minimum of 10km distance was maintained between sites to ensure the uniqueness of populations. Sites were identified with the help of herbarium records, land care groups and farming communities. The sites comprised roadsides, native and unmodified pastures, and protected areas such as reserves and box gum grassy woodlands. From each population 3 -5 seed heads each were collected from about 15 – 25 mother plants depending on the availability of the plants in a particular population.

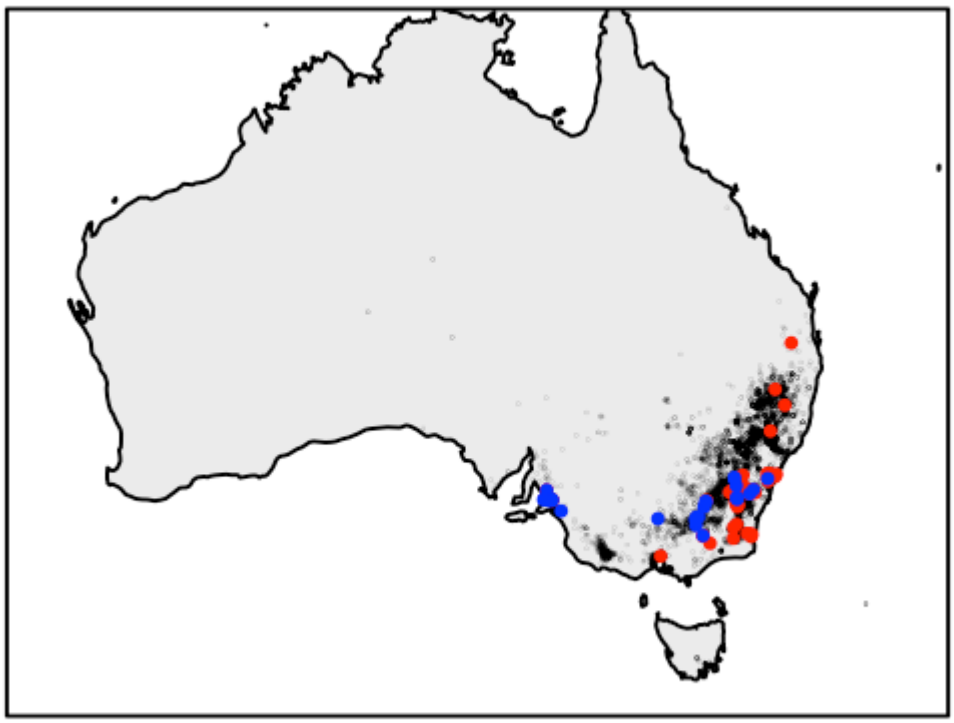


Figure 4 - 1. Distribution of populations used in different analyses of the study. The 44 seed sampling sites where seeds were collected for the common garden experiments in *B. macra* are represented in red dots. The 25 leaf sampling sites where leaves were collected for the genomic data analysis in *B. macra* are represented in blue dots. The species distribution of *B. macra* as recorded in Atlas of Living Australia is depicted as black dots.

For the purpose of conducting a genetic population structure analysis of *B. macra*, we collected leaf samples from 25 populations across the states of Victoria, NSW, and South Australia. A minimum of 10km distance was maintained between sites to ensure the uniqueness of populations. From each population leaf samples of about 15 individuals were collected and were placed in envelopes with silica beads, to dry the leaves upon collection. Out of the 25 populations used for the genotypic analysis, 11 populations have data common for both the phenotypic analysis and the genotype analysis conducted in the study.

Growth chamber experiments

The experiment was carried out in two controlled temperature rooms. Lights of the control temperature rooms were set to a 12-hour day/ night cycle, where 8 am to 8 pm was the day, while 8 pm to 8 am was the night, each set to a different temperature. The temperatures used for the control temperature treatment was 23°C from 8 am to 8 pm and 10° C from 8 pm to 8 am. Likewise for the increased temperature treatment was 28°C from 8 am to 8 pm and 13°C from 8 pm to 8 am. To decide on these temperature values, first temperature data belonging to sampling sites for current and predicted conditions (2070) were extracted from WorldClim database using R. As the germination of the species occurs in spring, the average highest temperature (average of the temperatures for month of November for all the locations) of the month of November (warmest month in spring) was used as the daytime temperature and the average lowest temperature of November was used as nighttime temperature. The average values based on current data were used as the control conditions while average values obtained from the predicted data were used to set the high temperature treatment values.

We used seeds collected from 44 populations for this experiment. Three to eight maternal plants each (mean = 5, range = 3-8) (Table 4 - S1) and ten seeds per maternal plant per population per treatment were used for this experiment. The availability of maternal plants and number of

viable seeds per maternal plant was different for each population. Seeds were surface sterilized to avoid fungal contamination by washing with distilled water, 70% ethanol and with 10% commercial chlorine bleach solution. The seeds were then patted dry with tissue paper before transferring them into the agar plates (55mm X 15mm) for germination. The germination medium used consisted of 1% agar and 0.1 % PPM (Plant Preservative Mixture) dissolved in distilled water. We randomly assigned the seeds from each maternal plant to two agar plates and assigned the plates to one of two treatments. Seeds were observed daily for 14 consecutive days for signs of germination (Nik et al. 2011; Nizam 2011). Seeds were considered to be germinated when the radicle had emerged about 2mm.

After 14 days of germination, the number of germinated seeds were counted to identify the proportion of germinated seeds for each population, all seedlings were removed from the agar and the length of the root and the shoot of each was measured using a digital calliper. Root and shoot lengths were measured at the individual seedling level for each mother plant in each population. After measuring the root and shoot lengths of seedlings, each seedling was divided into shoots and the roots. All roots and shoot parts belonging to individuals sourced from one petri plate (replicates of each maternal plant belonging to a population) were combined and put in two separate envelopes marked as roots and shoots. The envelopes were dried at 70 $^{\circ}$C in an oven for 48 hours until a constant weight was obtained. Then the dry biomass of each of the combined root and shoot tissues of each maternal plant was weighed using an electric balance. The average root biomass and the average shoot biomass for each maternal plant was calculated by dividing the mass by the number of seedlings included in each envelope.

DNA extraction, preparation of GBS libraries and sequencing

Genomic DNA was extracted using 30-40 mg of dried leaf tissues from the 251 samples using the DNAeasy Plant Mini Kit (Qiagen, Valencia, CA, USA). Quality of extracted DNA was checked by running samples on 2% agarose gel stained with ethidium bromide and viewing it under UV-Vis Spectrometer (NanoDrop 1000, Thermo Scientific). DNA quantity was assessed using a QuBit broad-sensitivity DNA quantification system (Invitrogen, Carlsbad, CA, USA). GBS libraries were prepared by following a modified version of the protocol from Poland et al., 2012. This involves a double digest (with *HF-PstI* and *MspI* restriction enzymes), genotype by sequencing (ddGBS) for equal amounts of DNA per individual. To modify the protocol, PCR

reactions were performed for individual samples and equal amounts of PCR products were pooled. Fragments between 400-600 bp were selected and sequencing of the libraries was done by GENEWIZ^{TM} (Suzhou, China) on Illumina HiSeq 4000 platform (paired end, 150 bp).

SNP calling and filtering.

Raw reads were demultiplexed and cleaned by filtering for low quality reads using STACKS (2.2) process_radtags (Catchen et al., 2013) with default settings. Then the reads were aligned to the *B.decipiens* reference genome (Chapter 2) using Burrows-Wheeler Aligner (BWA) (Li & Durbin, 2009). Samtools v 1.9 (Li et al., 2009) was used to transform the SAM files to BAM files for use within STACKS. The BAM files were filtered only to keep reads with a map quality greater than 30. The argument gstacks (using the default model "Marukilow") and populations were used in that order on the BAM files to create a VCF file with 1,472,554 unfiltered SNPs. Variant filtering was done as follows using VCFtools (Danecek et al., 2011). We removed individuals with reads less than 25000 reads, removed indels, filtered individual genotypes to have a depth above 5, removed variants genotyped in less than 50% of individuals, removed genotypes with a quality less than 20, filtered individuals with more than 50% missing data, filtered variants with a minor allele frequency less than 0.05 and selected only biallelic SNPs. The resulting SNPs were filtered to remove duplicate loci using the R\HDplot package (McKinney et al., 2017). This species is likely an allotetraploid (Sumadijaya, 2015) and we sought to target diploidized regions of the genome for downstream population structure analysis. According to the recommendations, loci with a *H* value (proportion of heterozygotes) between 0-0.6 and a *D* value (read ratio deviation) between (-7) - 7 were selected after plotting the results from running the HDplot function (Figure 4 - S1). The selected diploid loci were filtered for linkage using a single SNP per 1 kb window, resulting in 5,566 SNPs from 162 individuals (Table 4 - S2).

Climate variables

To test the effect of a range of climatic variables on the measured phenotypic variables, for each sampling location 19 bioclimatic variables (Booth 2022) were downloaded from WorldClim database (Fick & Hijmans, 2017) (Table 4 -S3) at 2.5-minute resolution using the latitude and longitude coordinates using the R\raster package in R (Hijmans & van Etten, 2012). Highly

correlated variables (Pearson's correlation coefficient > 0.70) were dropped by randomly removing one of each pair of highly correlated bioclimatic variables (Figure 4 - S2). This process reduced climate variables from 19 to seven (temperature annual range (var7), mean temperature of the driest quarter (var9), mean temperature of the coldest quarter (var11), precipitation seasonality (var15), precipitation of the driest quarter (var17), precipitation of the warmest quarter (var18) and precipitation of the coldest quarter (var19)). A principal component analysis was then done to summarize variation in climate among populations (Figure 4 - 2).

Trait analysis

A summary of all the trait responses measured are given in Table 4 - 1 below. All statistical analyses were conducted in R version 3.6.2. We tested for the normality of data within treatments using the Shapiro test in R. We improved normality of the data by log and square root transformation of the data where appropriate.

Table 4 - 1 *Bothriochloa macra* traits included in this study.

Trait response	**Description**
Root length	Length of the longest root emerged in a seedling
Shoot length	Length of the longest shoot emerged in a seedling
Root to shoot ratio	Ratio obtained by dividing the root length by the shoot length for each seedling
Root biomass	Mean dry biomass of all the roots from seedlings derived from a single mother plant

Shoot biomass	Mean dry biomass of all the shoots from seedlings derived from a single mother plant
Total biomass	Mean total biomass (sum of shoot biomass and the root biomass from seedlings derived from a single mother plant)
Proportion of germinated seeds	The number of seeds that produced radicals and grew into seedlings divided by the total number of seeds germinated per mother plant

Highly correlated measured traits (Pearson's correlation coefficient >0.70) were identified using a correlation matrix (Table 4 - S4). The least correlated trait measures were total biomass, root to shoot ratio and proportion of germinated seeds (Table 4 - S4).

A multivariate analysis of variance was done to identify if there were significant associations of the least correlated trait responses with the treatment and the first three climate PCs (PCclims). We also included the interactions between treatment and the PCclims but removing non-significant interactions using backwards elimination. Using the MANOVA we calculated the approximate F-statistics and Wilk's λ (multivariate F value) to measure the strength of the associations between the traits and the treatment as well as the first three climate PCs. To further explore the univariate relationships among all the measured traits and the PCclim variables, linear regression models were fit. We included PC1clim, PC2clim and PC3clim as well as treatment and all interactions with treatment and the climate PCs as fixed effects. In order to select the best fit models that explained the relationships among the traits and the different climate PCs, R\ 'MuMIn' (Bartoń, 2020) was used for model selection. The model with lowest AICc and AICc delta values was chosen as the best fit model for each linear regression. If the best fit model did not include the treatment effect, the next best fit model that included the treatment effect and also had a AICc delta value between 0-2 was used to explain any potential relationships. We retained treatment in the model as it was part of our original experimental design. AICc delta

values of 0-2 provide substantial support for model fit (Burnham & Anderson, 2007). Summaries of the model selection parameters for the linear models built to explain the relationship between traits and first three PCclims representing climate are given in Tables 4 - S5-S11.

To assess the amount of genetic variation for quantitative traits, we implemented mixed effect models using R package 'lme4' (Bates et al., 2014) for root length and shoot length as we had multiple offspring measured per mother for these traits. The data was structured in a nested design, where root and shoot lengths of each replicate (mean = 7, range = 2-15) belonging to each mother plant (mean= 5, range = 3-7) were nested under the population each mother plant belonged to. Population and mother were included as random effects and the treatments were analyzed separately. This allows us to estimate the total phenotypic variance attributed to each putatively half-sib family. Variance caused by maternal family effects (σ^2_f) and the residual variance (σ^2_r) were extracted using the VarCorr function. As the variance due to half- sib family (σ^2_f) is the product of additive genetic variance (σ^2_a) and coefficient of relationship between individuals within a family (r_{xy}) ($\sigma^2_f = \sigma^2_a * r_{xy}$) and r_{xy} for half sib family is 1/4 (Fisher, 1919), the additive genetic variance was calculated as $4* \sigma^2_f$. Using the value of additive genetic variance, narrow sense heritability (H_{ns}) was calculated. Narrow sense heritability is the proportion of phenotypic variance attributed to the variation in the additive effect of genes and can be calculated as; $H_{ns} = \sigma^2_a / (\sigma^2_a + \sigma^2_r)$ (Falconer & MacKay, 1996).

Gradient forest analysis of traits

A gradient forest analysis using the R\gradientforest package (Ellis et al., 2012) was used to determine which environmental variables from the 19 bioclimatic variables best explained trait turnover of the measured early life history traits of *B. macra*. Separate analyses were done for each treatment and population means of the traits in each treatment were utilized in the analysis. Gradient forest is a regression tree-based machine learning algorithm that searches for breaks across continuous environmental predictors. The model was run with 500 trees and a correlation threshold of 0.5. To visualise the gradient forest predictions across the range of the populations where the individual samples were collected, bioclimatic variables for 100,000 randomly generated geographic coordinates were extracted. These climatic variables were transformed for each random point using the gradient forest model. These values were summarised using a

principal component analysis (PCA). As suggested in the gradient forest manual, the top three principal components were transformed for visualisation, where the red value is defined by PC1+PC2, the green value by –PC2, and the blue by PC3+PC2-PC1.

Population genetic structure

Admixture (Alexander et al., 2009) was used to infer the population genetic structure using the filtered SNP set. Admixture was run with a with a major termination criterion of 1×10^{-9}, 1,000 bootstraps and 10-fold cross-validation for $K = 2–12$, where K equals the number of genetic groups. The K that produced the lowest cross-validation error was selected as the best K value. Admixture results were visualized using R\Pophelper 2.3.1 package (Francis, 2017) and pie charts. To further complement the population structure analysis a principal component analysis (PCA) was done using the R\SNPRelate package (Zheng et al., 2012). R\Poppr package (Kamvar et al., 2014) was used to detect if any clonal genotypes were present in the samples.

Isolation by distance and environment

Filtered SNPs were used to calculate pairwise F_{st} across all populations using the R\DartR package (Gruber et al., 2018). Pairwise geographic distance between populations was calculated using the R\geosphere package (Hijmans et al., 2017) from longitude and latitude coordinates. Pairwise correlations between the 19 bioclimatic variables representing the sampling locations used to collect individuals to generate SNP data and used to calculate F_{st} above were calculated and one of each pair of highly correlated variables (Pearson's correlation coefficient > 0.70) were removed (Figure 4 - S3). This reduced the bioclimatic variables from 19 to seven (mean diurnal range, var2; isothermality, var3; minimum temperature of warmest month, var6; mean temperature of the wettest quarter, var8; mean temperature of the warmest quarter, var10; precipitation of the warmest quarter, var18; and precipitation of the coldest quarter, var19). These seven environment variables were first scaled and cantered to account for the differences in magnitude. Then a measure of environmental distance between each population pair was calculated by pairwise Euclidean differences between the sites. A Mantel test was used to test the associations between linearized F_{st} calculated as (F_{st} / (1- F_{st}) with geographic distance and environmental distance using the R\ade4 package based on 999 permutations (Dray & Dufour, 2007).

Genetic diversity

For each population we calculated the standard measures of genetic diversity - expected heterozygosity (H_E), observed heterozygosity (H_o) and allelic richness (A_R) as average values across loci using the filtered SNP set. The 95% confidence intervals of A_R were calculated with 1000 bootstraps. These were calculated using the R\diveRsity package (Keenan et al., 2013).

F_{st}-Q_{st} analysis

To compare the Q_{st} of the two quantitative traits root length and the shoot length used in the mixed models above (from the control treatment) to the mean F_{st} of series of marker loci, the R function QstFstComp (Gilbert & Whitlock, 2015) was used. This function calculates the distribution of Q_{st}-F_{st} under a model assuming neutrality of both the phenotypic trait and the genetic markers from which F_{st} is estimated. Eleven populations that had both trait data and SNP data were selected for this analysis. The calculations were done after setting the "breeding.design" option to half-sib dam model, which assumes that dams are nested within populations with all offspring coming from separate sires. Calculations were based on 10,000 simulations.

Results

Climate variation

The first three PCclims explained 82.40% of the total variation (PC1clim-35.8%, PC2clim- 25.9% and PC3clim-20.7%) (Figure 4 - 2). The relationships between the first three climate PCclims and the original climate variables are illustrated in Table 4 - 2. PC1clim was positively associated with mean temperature of the coldest quarter and precipitation seasonality and negatively with temperature annual range and precipitation of the driest quarter while PC2clim was positively associated with precipitation of the warmest quarter and precipitation of the driest quarter, and negatively associated with mean temperature of the driest quarter. Finally, PC3clim was negatively associated with precipitation of the coldest quarter and positively with mean temperature of the driest quarter.

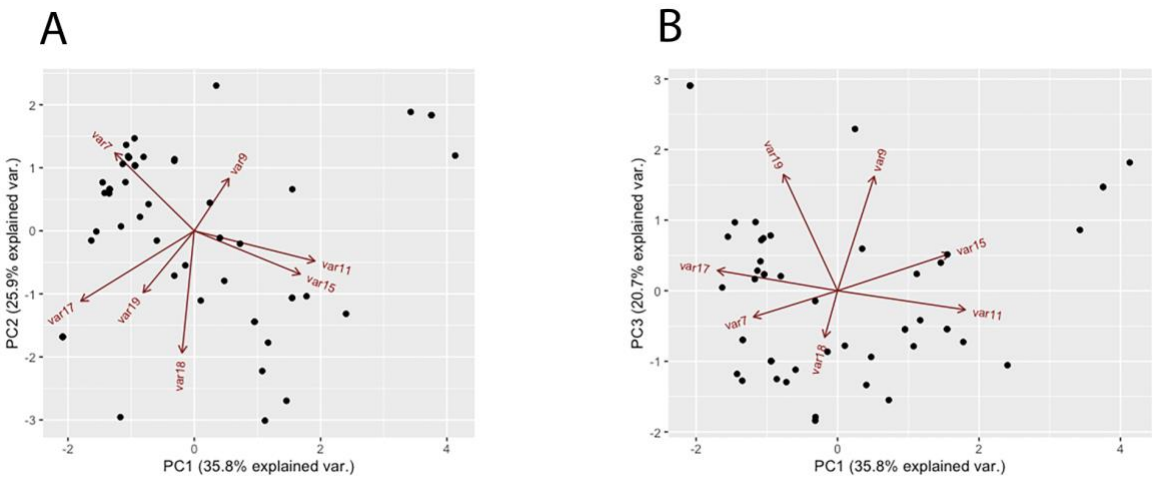


Figure 4 - 2. Biplots of the principal component analysis for the climatic variables used in the analysis of trait variation. (A) PC1clim vs PC2clim. (B) PC1clim vs PC3clim. The percentage explained by the primary principal components is shown on each axis. The climatic variables represented by each of the abbreviations is given in Table 4 - S3.

Table 4 - 2. The correlation matrix between the first three PCclims and the original climate variables that represent them. Correlations greater than 0.45 used to describe the relationships between the original climate variables and PCclims are given in bold text.

	PC1clim	**PC2clim**	**PC3clim**
Temperature annual range (var7)	**-0.6526384**	-0.25985945	0.323833034
Mean temperature of driest quarter (var9)	0.0732246	**-0.48013598**	**0.747161994**
Mean temperature of coldest quarter (var11)	**0.8744716**	0.05839364	-0.007957052

Precipitation seasonality (var15)	**0.7868233**	0.15085036	-0.200747012
Precipitation of the driest quarter (var17)	**-0.5730069**	**0.72300223**	-0.301205699
Precipitation of the warmest quarter (var18)	0.1977918	**0.89150523**	0.247151296
Precipitation of the coldest quarter (var19)	-0.2371788	0.27894392	**-0.883350195**

Trait variation

The MANOVA was done to identify if multivariate trait variation was associated with the treatment and the first three climate PCs. We found that the traits had significant associations with treatment, PC2clim and PC3clim but not PC1clim (Table 4 - 3). Univariate analysis found that treatment had a significant effect in five of the seven traits (root length, shoot length, root biomass, shoot biomass and total biomass). In all of these cases the warm treatment increased the traits (Table 4 – 4(a)). Only two traits, shoot length and germination proportion were associated with PC1clim, however, the overall model fit was not significant for germination proportion. Shoot length was positively associated with PC1clim indicating greater shoot length in populations experiencing warmer temperatures in the coldest quarter and greater precipitation seasonality. Shoot biomass and total biomass were significantly associated with PC2clim, revealing greater biomass of populations in regions with more precipitation in the driest quarter and warmest quarter. Similarly, these traits were negatively associated with PC3clim, where plants with greater biomass were found in regions with lower temperatures in the driest month and higher precipitation in the coldest quarter. Root to shoot length ratio was also negatively associated with PC3clim but the overall for the model R^2 was low (0.06) (Table 4 - 4(a)(b)). In all cases the interaction effect was not significant and there was only a main effect of treatment, suggesting a similar response of populations to the temperature treatment regardless of the

climate of origin.

Table 4 - 3. A summary of the MANOVA of root to shoot ratio, total biomass and proportion of seed germination population mean trait responses of *Bothriochloa macra* to treatment and the first three principal components representing climate variation. We report the approximate F statistic with degrees of freedom as subscript and symbols specifying significance of effect, in addition to Wilk's λ (multivariate F value).

	F statistic with degrees of freedom	**Wilk's λ**
Treatment	19.8881, 79 ***	λ = 0.36203
PC1clim	1.1566, 79	λ=0.90704
PC2clim	2.4928, 79*	λ=0.81908
PC3clim	2.2383, 79*	λ=0.83449

*p<0.05, **p<0.01, ***p<0.001

Table 4 - 4. Results of best fit general linear models for the measured *Bothriochola macra* traits. The relationship between the population means of the traits (response) and treatment as well as the first three PCclims. a) The slope estimates and significance (bolded p values are significant at the 0.05 level). b) Summary of the parameters explaining the overall model fit (bolded p values are significant at the 0.05 level).

a)

Trait	**Variable**	**Estimate**	**Standard error**	**t value**	**p value**
Root length	Treatment	1.0808	0.3892	2.777	**6.73×10^{-3}**
	PC3clim	-0.3143	0.1613	-1.949	5.45×10^{-2}
Shoot length	Treatment	1.3411	0.1520	8.821	**1.04×10^{-13}**
	PC1clim	0.1055	0.4699	2.244	**2.73×10^{-2}**
Root to shoot ratio	Treatment	-0.1682	0.1293	-1.300	1.97×10^{-1}
	PC3clim	-0.1262	0.0536	-2.356	**2.07×10^{-2}**
Root biomass	Treatment	0.5026	0.1187	4.234	**5.67×10^{-5}**
	PC3clim	-0.0700	0.0491	-1.424	1.58×10^{-1}

Shoot biomass	Treatment	0.1936	0.0259	7.458	**6.52×10^{-11}**
	PC2clim	0.0362	0.0099	3.632	**4.78×10^{-4}**
	PC3clim	-0.0221	0.0108	-2.046	**4.38×10^{-2}**
Total biomass	Treatment	0.4960	0.0707	7.014	**4.96×10^{-10}**
	PC2clim	0.0777	0.0271	2.863	**5.27×10^{-3}**
	PC3clim	-0.0620	0.0295	-2.104	**3.82×10^{-2}**
Proportion of germinated seeds	Treatment	-0.0143	0.0430	-0.333	7.39×10^{-1}
	PC1clim	0.0292	0.0132	2.198	**3.06×10^{-2}**

b)

Linear model	**Degrees of freedom**	**Adjusted R^2 value**	**F statistic**	**Model p value**
Root length ~ Treatment +	87	0.10	5.75	**4.5×10^{-3}**

PC3clim				
Shoot length ~ Treatment + PC1clim	87	0.48	41.43	**2.29×10^{-13}**
Root to shoot ratio ~ Treatment + PC3clim	87	0.06	3.62	**3.08×10^{-2}**
Root biomass ~ Treatment + PC3clim	87	0.17	9.97	**1.25×10^{-4}**
Shoot biomass ~ Treatment + PC2clim +PC3clim	86	0.43	23.83	**2.57×10^{-11}**
Total biomass ~ Treatment + PC2clim + PC3clim	86	0.39	20.19	**5.38×10^{-10}**
Proportion of germinated seeds ~ Treatment + PC1clim	87	0.03	2.47	9.04×10^{-2}

Heritability of root and shoot length.

Estimating the proportion of variation caused by genetic variance can give insight into the short-term capacity of traits to respond to environmental change. In the mixed effect model that modelled the variation of root length in the control treatment with both maternal family and population random effects, 50% of the variation was caused by the maternal family random effect, while 43% of the variation was due to population random effect (Table 4 – 5(a)).

Consequently, the total amount of trait variation caused by maternal (genetic and maternal factors) or genetic differentiation among populations was 93%. When we assumed unrelated sires and negligible maternal effects, the calculated value for narrow sense heritability for root length was 0.96. In the mixed effect model that modelled the variation of shoot length in the control treatment with both maternal family and population random effects, 50% of the variation was caused by maternal family random effects, while 28% of the variation was due to the population random effect (Table 4 – 5(b)). Consequently, the total amount of trait variation caused by maternal (genetic and maternal) factors was 78%. Given the above assumptions, for shoot length, the value for narrow sense heritability was estimated to be 0.89.

Table 4 - 5. Results of the mixed effects models for *Bothriochloa macra* traits (root length and shoot length) in the control treatment. The variation of traits due to maternal and population random effects. a) Root length variation with random effects. b) Shoot length variation with random effects.

a)

Random effect	**Variance**	**Percentage variance (%)**	**Standard deviation**
maternal	121.20	49.6	11.009
population	105.95	43.3	10.293
Residual	17.34	7	4.165

b)

Random effect	Variance	Percentage variance (%)	Standard deviation
maternal	17.617	48.9	4.197
population	9.968	27.7	3.157
Residual	8.459	23.5	2.908

Gradient forest analysis

Based on the gradient forest analysis using all the traits measured under control treatment as response variables, out of the 19 climate variables var3 (isothermality), var15 (precipitation seasonality), var19 (precipitation of the coldest quarter), var4 (temperature seasonality) and var14 (precipitation of the driest month) were identified as the most important climate variables that explained trait turnover across populations (Figure 4 - 3(A)). Out of all the measured traits used in this analysis only four traits were used to build the gradient forest model. They were shoot biomass, root to shoot ratio, total biomass, and proportion of germinated seeds. These traits were the ones with positive R^2 values in the model, in that those traits could be predicted by the available climate variable predictors. Differences in environmentally associated phenotypic variation can be seen across the range (Figure 4 - 4(A)). A striking red region in the centre of the range was apparent suggesting unique trait composition in this region. Biplots of individual traits used in the model against the topmost important variables, suggested this pattern was in part explained by reduced germination (majority of populations with lower values but overall there is a variability in the rates) in a region with low precipitation seasonality and isothermality (Figure 4 - S4).

Based on the gradient forest analysis using the all the traits measured under temperature treatment, out of the 19 climate variables var15 (precipitation seasonality), var16 (precipitation of the wettest

quarter), var17 (precipitation of the driest quarter), var13 (precipitation of the wettest month) and var19 (precipitation of the coldest quarter) were the most important climate variable that explained turnover in traits among populations (Figure 4 - 3(B)). Out of all the measured traits used in this analysis only three traits were used to build the gradient forest model. They were shoot length, shoot biomass and total biomass. These were the traits with positive R^2 values in the model, where those traits could be predicted by the available climate variable predictors. Differences in environmentally associated phenotypic variation can be seen across the range (Figure 4 - 4(C)). As above, a distinct region in the centre of the range was identified and biplots suggested that the pattern was in part explained by reduced shoot length and biomass (majority of populations with lower values but overall there is a variability in the values recorded) in regions experiencing reduced precipitation seasonality and precipitation in the wettest quarter (Figure 4 - S5).

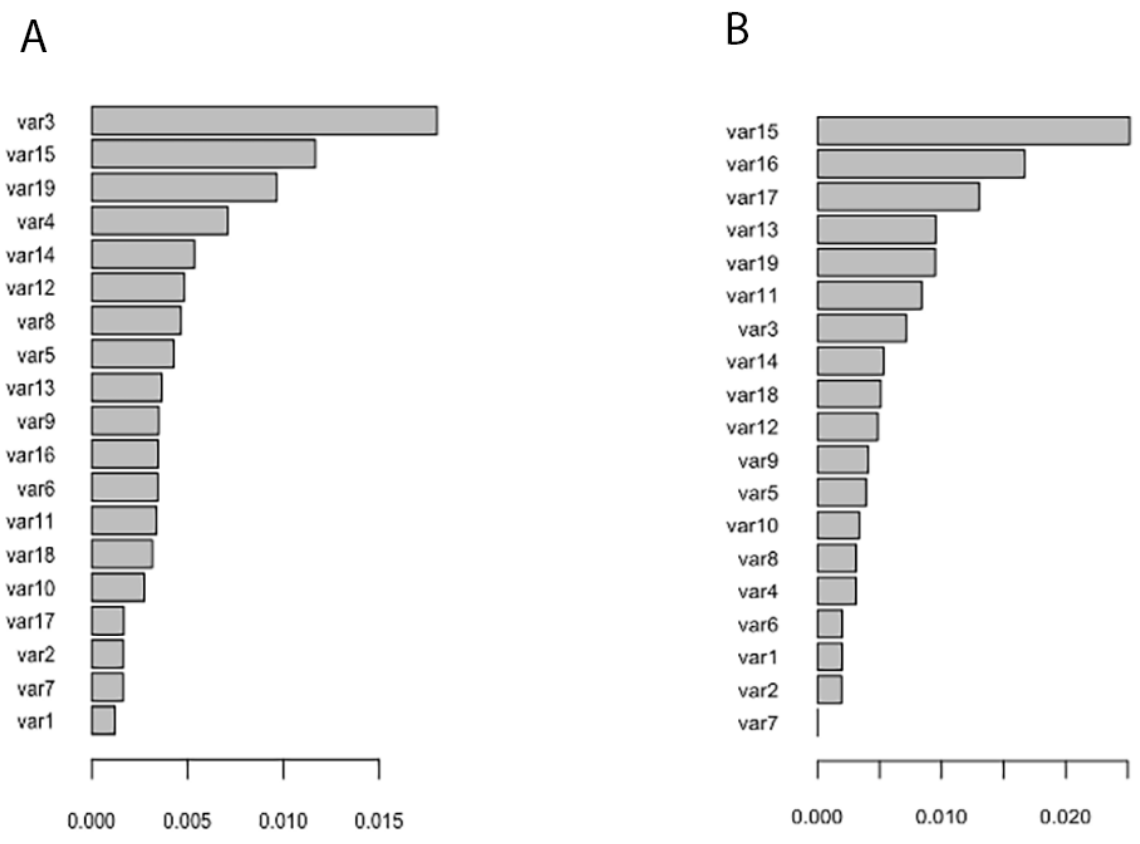


Figure 4 - 3. Ranked importance plots of the gradient forest models. (A). Ranked importance of the 19 environmental variables on the phenotypic variation in the traits measured in control

treatment of the populations based on the gradient forest analysis. (B). Ranked importance of the 19 bioclim variables on the phenotypic variation in the traits measured in temperature treatment of the populations based on the gradient forest analysis.

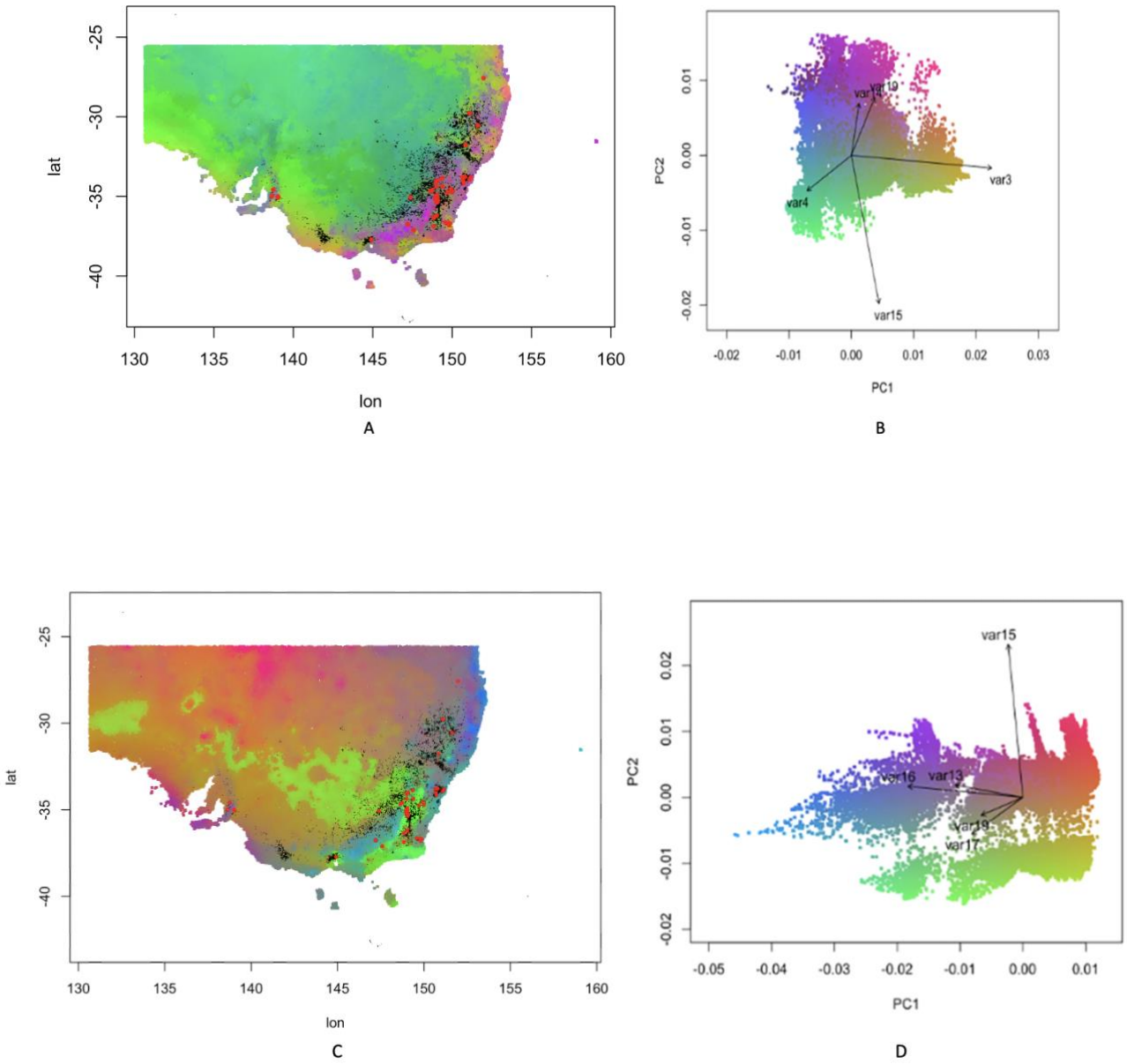


Figure 4 - 4. Visualization of the gradient forest model results. (A) Gradient forest transformed climate variables plotted on a map using the randomly generated geographic coordinates using the model run with all measured control treatment traits. Similar colours reflect similar phenotypic variation. Points on map reflect sampled locations. (B) Principal components analysis (PCA) of transformed climate variables which represent the colours of the map in (A). (C) Gradient forest transformed climate variables plotted on a map using the randomly generated geographic coordinates using the model run with all traits measured in the temperature treatment. Similar

colours reflect similar phenotypic variation. Points on map reflect sampled locations. (D) Principal components analysis (PCA) of transformed climate variables which represent the colours of the map in (C).

Population genetic structure

No clonal genotypes were detected among the loci used for the population structure analysis. In the Admixture run the K that produced the lowest cross-validation error was k=8 (Figure 4 - 5(A), Figure 4 - S6). Individuals that belonged to the same population tended to be grouped into a subset of genetic clusters, and there appeared to be some geographic clustering by region (Figure 4 - 5). In populations from the west (mostly South Australian) the dark blue cluster was much more common compared to the eastern populations, while the lime green cluster was more common in eastern populations. Populations in the centre of the distribution shared a light blue cluster that is not common among the populations from the far east or far west. The yellow cluster seemed to be shared among populations throughout the distribution in varying amounts.

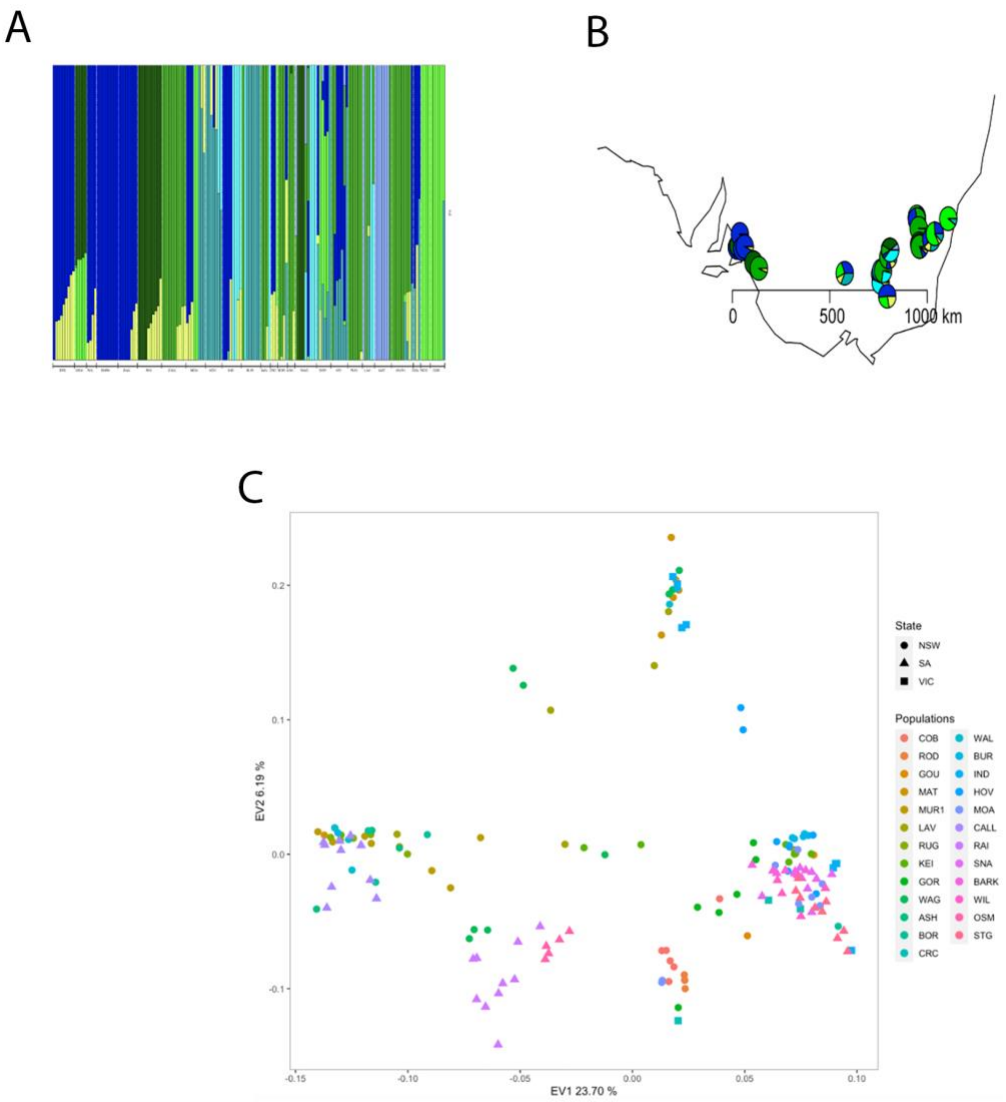


Figure 4 - 5. Genetic differentiation of sampled populations of *B. macra*. (A) A distruct plot showing the results of Admixture run of the filtered SNP set for k=8. Populations from West to East are arranged left to right in the plot. (B) Population pie charts for k=8 mapped to represent their geographical distribution, admixture proportions for each population are displayed, colours correspond to the clusters in the distruct plot. (C) Principal component analysis of the filtered SNP set, first two eigenvectors are presented, individuals are coloured according to their population.

Isolation by distance and environment

Linearized pairwise genetic distance between locations (F_{st} / (1- F_{st}) was significantly correlated with geographic distance, suggesting signals of isolation by distance (Mantel's r = 0.20, *P*-value = 0.01). Linearized pairwise genetic distance between locations (F_{st} / (1- F_{st}) was also correlated with environment distance, suggesting signals of isolation by environment (Mantel's r = 0.18, *P*-value = 0.03).

Genetic diversity

Average values of calculated standard measures of genetic diversity were as follows- H_E - 0.07 (range 0.04-0.12), H_o - 0.08 (range 0.06-0.14) and A_R - 0.96 (range 0.94-1.12). Calculated values for each population can be found in Table 4 - S12.

Q_{st} - F_{st} analysis

For both root length and shoot length measured under control treatment the calculated Q_{st} value was less than the calculated F_{st} value indicating that quantitative trait genetic differentiation among populations is more likely due to stabilizing selection (Table 4 - 6).

Table 4 - 6. Summary of the Q_{st}-F_{st} analysis.

Quantitative trait	**Q_{st}-F_{st} [95% CI]**	**F_{st} [95% CI]**	**Q_{st} [95% CI]**
Root length	-0.26 [-0.20, 0.20]	0.32 [0.31, 0.32]	0.05 [0.0003, 0.14]
Shoot length	-0.27 [-0.20, 0.21]	0.32 [0.31, 0.32]	0.04 [-0.003, 0.12]

Discussion

Using trait-environment associations measured in common gardens, we found evidence consistent with local adaptation to climate in the *B. macra.* Specifically, we found that in seedlings, biomass related traits were greater in populations experiencing more precipitation in warm and dry times of year. By contrast, the reduced growth in arid regions is suggestive of the evolution of a drought tolerance strategy in these *B. macra* populations. However, other explanations are possible, including the role of maternal effects, or neutral divergence. Indeed, our analysis of population structure identified that genetic distance was associated with both geographic and environmental distance. Further, analysis of divergent selection on the seedling traits of root and shoot length using a Q_{st}-F_{st} did not yield evidence of local adaptation and instead was consistent with stabilizing selection. Still these data raise the possibility that climate is important in structuring trait variation in this species, and thus should be considered when selecting seeds for restoration efforts.

Climatic patterns in seedling traits

The most salient pattern of our trait-climate analysis was that biomass increased in high rainfall and low temperature areas. This pattern was found irrespective of the temperature treatment. Specifically, shoot and total biomass showed associations with both PC2clim and PC3clim. PC2clim, which was positively correlated with both traits, was also strongly positively correlated with precipitation of the warmest and the driest quarters (Table 4 – 4(a), Table 4 - 2). PC3clim, which was negatively associated with both traits, was strongly negatively correlated with precipitation of the coldest quarter, and positively correlated with temperature of driest quarter (Table 4 – 4(a), Table 4 - 2). Consequently, the predominant pattern was for seedlings with reduced biomass in drier and warmer regions. Low rainfall and high temperature combinations can cause a decrease in soil moisture resulting in drought conditions (Assani et al., 2016; Gergis & Ashcroft, 2013; Kirono et al., 2017). One of the many definitions of drought is related to soil moisture, emphasizing the availability of soil moisture to support vegetation growth (Kiem et al., 2016). Depletion of soil moisture can be a result of lack of rainfall and/or excess evaporation of existing soil water, which can be caused by high temperatures. Plants display numerous drought avoidance and tolerance strategies. Some show increased development in the root system to maintain productivity of water uptake during times of scarcity

(Comas et al. 2013). Reduction in root and shoot biomass in seedlings could be expected as plants increase the efficiency of nutrient utilization than the efficiency of nutrient uptake during water deficit conditions (Hu et al. 2006; Mahmood et al. 2022). Overall, our observation of reduced total biomass in more arid regions is consistent with the evolution of a drought tolerance strategy in *B. macra* seedlings. Our results corroborate a meta-analysis done by (Roybal & Butterfield, 2018) on functional trait heritability and climatic adaptation among grasses. They found that total biomass increased with increasing annual and dry season precipitation indicating that precipitation variables were frequent drivers of local adaptation among species of grasses. In all cases the interaction effect was not significant and there was only a main effect of treatment, suggesting a similar response of populations to the temperature treatment regardless of the climate of origin. This suggests that any plastic response was consistent along the climate gradients.

The Gradient Forest analysis largely supported the findings of the linear models, however, an advantage with this approach is that it allows for threshold effects and nonlinear relationships between the traits and the environmental variables (Ellis et al., 2012). In the control treatment, var15 and var19 were among the top three important variables in explaining the trait turn over across the landscape (Figure 4 - 3(A)). The regression trees were built using the traits shoot biomass, root to shoot ratio, total biomass, and proportion of germinated seeds. Similarly, var15 was strongly associated with PC1clim and var19 was strongly associated with PC3clim and either PC1clim or PC3clim were retained in the best fit model for all of these traits (Table 4 – 4(a)). A somewhat similar pattern of trait turnover across the landscape was also identified when the temperature treatment was used (Figure 4 - 4(C)), although there were some differences in the importance plots (Figure 4 - 3(B)) and traits retained in the model. In both treatments there appears to be a similar geographic pattern across the landscape in trait turnover with respect to climate which may be potentially leveraged to identify seed zones for restoration using a clustering approach. This is especially the case if these geographic patterns extend to other traits or candidate loci for climate adaptation. For instance, regions in the centre of the range appeared to have divergent trait patterns in both treatments (Figure 4 - S4, 4 - S5). Further this random forest approach can be used to predict regions more at risk of future maladaptation from climate change and we intend to extend this approach using both trait and genetic data in future studies.

Evidence for stabilizing selection on seedling traits

Shoot and root length, the traits for which we were able to calculate Q_{st}, did not show signals of local adaptation. Root length population means showed no association with historic climates (Table 4 – 4(a)), while shoot length showed a weak but significant association with PC1clim (Table 4 – 4(a)). Further, neither of these traits for the control treatment were retained in the Gradient Forest models, meaning that variation in these traits could not be predicted by the regression trees using the climate data. For the temperature treatment only shoot length was retained in the model. Importantly, although our estimates of heritability for both traits were high (root length = 0.96, shoot length = 0.89) using within population estimates of V_a, the results of the Q_{st}-F_{st} comparison did not indicate the presence of divergent selection. Instead, the pattern of quantitative trait genetic differentiation for these traits is more likely due to stabilizing selection ($Q_{st} < F_{st}$). Other studies have found evidence for stabilizing selection on traits related to drought tolerance despite gradients in precipitation even for traits known to directly influence drought tolerance, such as cavitation in trees (Lamy et al., 2011; Wortemann et al., 2011), perhaps because of constraints like pleiotropy. If root and shoot length are under climate mediated selection, fitness related traits such as growth generally respond slower to directional selection than non-fitness traits (Kingsolver et al., 2001; Merilä & Sheldon, 1999). Evolutionary change of fitness traits through directional selection may be prevented due to their more complex genetic background and genetic correlation among traits (Kruuk et al., 2008; Morrissey et al., 2012) This may also help to explain the lack of signals of divergent selection in the growth-related quantitative traits of root and shoot length. In contrast, in some grasses fitness related traits are under strong divergent selection (Razzaque and Juenger 2022; Rauschkolb et al. 2022).

The literature on the sources of bias in estimates of Q_{st}-F_{st} is extensive (Edelaar et al., 2011; Edelaar & Björklund, 2011; Miller et al., 2008; Pujol et al., 2008; Whitlock, 2008) and we will not review them all here. However, there are several prominent issues related to these estimates that should be mentioned. First, maternal effects can cause biases in these estimates. Estimation of Q_{st} can be biased downwards if maternal effects cause siblings to appear more similar and inflate the estimation for additive genetic variance (Gilbert & Whitlock, 2015). Maternal effects can also increase the variance among populations due to plasticity, biasing the estimate of Q_{st} upwards (Gilbert & Whitlock, 2015). We assumed families were composed of half siblings, but this may be incorrect and full siblings might be possible, especially if selfing is present. Further, selfing can cause complications when estimating Q_{st} due to effects like purging of deleterious alleles and heterozygote advantage after subsequent outcrossing (Gilbert & Whitlock, 2015).

Finally, although we used strict filters to remove duplicated loci, our estimates of molecular divergence may also be biased by the presence of undetected duplicated loci which could further reduce the reliability of this approach (Meirmans et al., 2018; Meirmans & Van Tienderen, 2013). Presumably, however, SNPs caused by the allopolyploidy would be shared across populations and bias F_{st} downwards meaning that this is unlikely to explain our finding of stabilizing selection on these traits.

We found that a large percentage of the root and shoot length variation was due to differences among maternal families (Table 4 - 5) and our estimates of narrow sense heritability were high. Although potentially confounded by maternal effects, and unknown relationships among siblings derived from open pollinated seed, our findings suggest substantial levels of heritable variation are likely for these traits. Heritability largely determines the magnitude and speed of phenotypic change in response to selection (Falconer & Mackay, 1996) therefore seedlings of *B. macra* have adaptive potential for these critical early life history traits, which may aid in responding to future environmental change.

Heritability measures reported in other studies investigating grasses are variable. Very high heritability was obtained for traits related to flowering while moderate heritability was observed for phenotypic traits such as panicle number and size (González Barrios et al. 2016). Moderate to very high heritabilities were reported for many phenotypic characters in several agriculturally important grasses (Janakiram et al. 2014). In a meta-analysis it was revealed that traits: leaf size, specific leaf area and total biomass in grasses showed very high heritability (Roybal and Butterfield 2018).

Population structure

Our population structure analysis identified the presence of strong population differentiation (Figure 4 - 5). Such population structure is common in other grass species (Ahrens et al. 2020; Gould et al. 2018; Guo et al. 2021; Hamann et al. 2016; Liu et al. 2016; Muktar et al. 2019). Population genetic structure and genetic diversity in grasses is influenced by their breeding system, with more genetic divergence and lower diversity among self-pollinating species than among outcrossing ones (Godt & Hamrick, 1998). Selfing and cleistogamy is known in the group (chapter 2), however *B .macra* seems to have varying proportions of chasmogamous and

cleistogamous flowers, with the majority being chasmogamous (Yu et al., 2000). However, we could not estimate population selfing rates using the genetic markers as the unintentional inclusion of paralogs could strongly bias the estimates. Presence of moderate but significant signals of IBD and IBE provided evidence for spatially and environmentally structured differentiation among populations (Prunier et al. 2017; Huang et al. 2016). Similarly, the Admixture plots did not reveal strong geographic structuring of the identified clusters, although there were general trends. For example, the blue cluster was more common in the west than the east, and the lime green cluster followed the opposite pattern. By contrast, the populations in the centre of the range appeared to be more admixed. This region also stood out in the Gradient Forest analysis suggesting that this region has a divergent trait profile. A further investigation of common garden traits of *B. macra* beyond the brief window of time investigated here would be critical in assessing the extent of trait differentiation across the landscape and how this might correlate with the genetic structuring we have identified.

Another concern could rise due to the presence of high population differentiation in this system. Epistatic incompatibilities and the break-up of coadapted gene complexes can cause a reduction in fitness during mixing of genotypes from highly differentiated populations (Schiffers et al. 2013). However, the lack of clear geographic division in structure clusters and admixture between clusters suggests that strong reproductive barriers among regions are not present. Presence of ploidy differences in this system could cause high population differentiation and pose a significant risk to fertility when cytotypes are mixed (Etterson et al. 2007; Bingham et al. 1994). We have not detected evidence of ploidy variation within *B. macra*, however it is closely related to *B. decipiens,* which is diploid, and they have overlapping geographical ranges. Hybridization within the clade has been documented (Sumadijaya 2015) and may be contributing to the unusual patterns of population structure we observed. Investigation of the extent of hybridization and the fitness consequences of both interspecific crosses and inter population crosses within a species are warranted.

Alternative explanations for the trait-climate patterns

Our common garden study demonstrated genetic differentiation in a key early life history trait

among populations. Because biomass was correlated with climate variables, our data are consistent with the evolution of a drought tolerance strategy in more arid environments. However, these patterns could also reflect genetic differences driven by adaptation to correlated environmental variables (Méndez-Toribio et al., 2020), or selection acting via an alternative mechanism (Nguyen-Queyrens & Bouchet-Lannat, 2003), or even neutral genetic changes. Such non-adaptive explanations should be considered given the evidence for isolation by environment that we uncovered using genome wide markers, many of which are likely neutral, the lack of evidence supporting divergent selection on root and shoot lengths using the Q_{st}-F_{st} approach, and the overlap between admixture in the centre of the range and striking patterns of trait turnover. Further phenotypic trait differentiation among populations can be generated by phenotypic plasticity in response to environmental cues, and this can occur across generations through transgenerational plasticity (Turcotte & Levine, 2016). This is especially important for early acting life history traits as measured here and this may offer an alternative explanation for the patterns we have found. Future investigation, including provenance trials where selection on traits could be quantified and local adaptation to drought measured under field conditions (Montwé et al., 2015) will be important in elucidating the cause of these patterns, particularly if the seeds used were derived from at least one generation of growth in common conditions to reduce transgenerational plasticity.

Implications for restoration

A substantial threat is posed to the conservation and management of ecosystems worldwide by the new evolutionary pressures caused by global climate change (Kumar et al., 2020). Drought is a serious natural disaster that has captured worldwide attention in recent decades. There has been an increase in the frequency and severity of extreme climate events (droughts, heat waves) in Australia recently (Cleverly et al., 2019; Ellis & Albrecht, 2017) and large areas of the country have been experiencing drought since 2017 (King et al., 2020). Unfortunately, future predictions about Australian drought conditions are dire. For example, the Southeast Australian Climate Initiative (SEACI) predicts a 9% decrease in rainfall in southeast Australia per 1 °C increase in annual global average temperature (Timbal, 2009). Studies have revealed that drought and heat waves will affect grasslands and their ecosystem services in particular worldwide (Schlaepfer et al., 2017; Sloat et al., 2018).

The increase in the duration and frequency of drought due to climate change is predicted to result in greater seedling mortality due to the increased stress on plants starting from the early growth period (Engelbrecht et al., 2006; Gilbert et al., 2001; Tobin et al., 1999). Early life stages are more susceptible to climate related mortality than later life stages (Haig et al., 1941; Leck et al., 2008). Consequently, traits facilitating drought resistance in provenances used for restoration of grasslands, particularly in early life stages that are key to establishment success, may be critical for long term success of restoration activities. Information gained from common garden studies such as this one, can aid the development of less conventional provenancing strategies that seek to adapt populations more quickly to future climate such as 'admixture provenancing' (Breed et al., 2013) or 'climate adjusted provenancing' (Prober et al., 2015). For example, Bustos-Salazar et al., 2017 found evidence that seedlings of *Drimys winteri* var *chilensis* obtained from xeric regions were most tolerant to drought than seedlings from other wetter parts and they recommend the use of seedlings from xeric regions for restoration of future drought prone areas. Similarly, seedlings of ponderosa pine grown from seeds obtained from drier regions survived better in experimental drought than other provenances, consistent local adaptation to drought (Kolb et al., 2016). Our data are suggestive of local adaptation of seedling traits to drought in *B. macra* implicating that sourcing seeds from drier areas as restoration stocks when restoring grasslands in areas prone to more drought in the future may be prudent. Extensions of the gradient forest approach using both genomic and trait data to predict provenances more suitable for future climates should be conducted and our analyses represent the first step in this process.

Conclusion

Our findings suggest the presence of local adaptation to climate related variables in the study species *B. macra*, where populations from more arid regions have evolved a 'drought tolerance' mechanism to survive drought by decreasing resource investment in overall growth and total biomass allocation. The detected signals of high population differentiation, and admixture among structure clusters, may be caused due to many factors related to the evolutionary history of the species, including mating system or even hybridization. Future studies investigating the extent of hybridization and fitness consequences of intra-population crosses are warranted, although our phenotypic data point to a benefit from a climate adjusted provenance strategy for this species.

Appendix III - Supplemental information for : Australian native grass *Bothriochloa macra* shows signals of local adaptation to drought.

Table 4 - S1. Summary of the location information and maternal plant number of each population used to obtain phenotypic data.

Population name	**State**	**Latitude**	**Longitude**	**Number of maternal plants used in phenotypic study/used in genotypic study**
COOM	NSW	-36.207	149.023	6
LAV*#	NSW	-34.912	149.025	5
SYD	NSW	-33.819	151.163	5
YAR	NSW	-35.301	149.095	4
BEM	NSW	-36.638	149.579	6
MAR	NSW	-34.707	150.007	4
WALA	NSW	-35.203	148.964	6

ARA	ACT	-35.259	149.085	3
BER	NSW	-36.356	148.844	3
KAM	ACT	-35.373	149.050	7
CLE	NSW	-33.927	151.097	7
BRT	VIC	-36.779	147.216	8
ARM	NSW	-30.531	151.617	6
MURR	NSW	-31.764	150.837	5
OXL	NSW	-33.773	150.800	3
SNA*#	SA	-35.009	139.015	8
STR	VIC	-37.719	144.907	6
CALL *#	SA	-35.067	139.048	7
CRC*#	VIC	-36.747	147.194	6

COR	NSW	-36.874	148.838	8
GRA	NSW	-34.000	149.085	4
MCK	QLD	-27.564	151.971	5
MAT*[#]	NSW	-34.975	149.020	6
OSM*[#]	SA	-34.970	138.653	5
OME	VIC	-37.112	147.576	6
FRO	NSW	-36.791	149.800	7
REI	NSW	-34.080	148.991	5
BIN1	NSW	-34.370	149.379	6
GOU*[#]	NSW	-34.754	149.714	7
ROD*[#]	NSW	-34.558	149.921	4
MUR1*[#]	NSW	-35.010	149.048	3

GIN	NSW	-35.011	149.047	5
COB*#	NSW	-34.031	150.691	5
BEG	NSW	-36.693	149.840	4
BUK	NSW	-36.679	149.787	7
MUL	NSW	-33.789	150.660	5
TUE	NSW	-33.866	149.352	4
DOU	NSW	-34.157	150.721	6
RUG*#	NSW	-34.370	148.995	7
KAN	NSW	-34.602	148.754	4
INV	NSW	-29.775	151.099	3
WAG2	NSW	-35.068	147.377	6
BINA	NSW	-34.672	148.628	5

JEI	NSW	-35.020	148.959	5

*Populations used in the Q_{st}-F_{st} comparison #Populations overlapping with genetic study

Table 4 - S2. Summary of population location information and the number of individuals for each population used in the population genomic analysis.

Population name	**State**	**Latitude**	**Longitude**	**Number of individuals**
ASH	NSW	-35.121	147.343	3
BARK	SA	-35.082	138.869	9
BOR	NSW	-35.375	147.252	4
BUR	NSW	-35.859	146.840	8
CALL*	SA	-35.067	139.048	10
COB*	NSW	-34.031	150.691	6
CRC*	VIC	-36.747	147.194	3
GOR	NSW	-33.973	148.887	6

GOU*	NSW	-34.754	149.714	3
HOV	NSW	-35.976	146.784	7
IND	VIC	-36.258	146.814	8
KEI	NSW	-34.057	148.891	7
LAV*	NSW	-34.912	149.025	5
MAT*	NSW	-34.975	149.020	7
MOA	NSW	-35.941	144.754	8
MUR1*	NSW	-35.010	149.048	9
OSM*	SA	-34.970	138.653	5
RAI	SA	-35.572	139.525	10
ROD*	NSW	-34.558	149.921	4
RUG*	NSW	-34.370	148.995	6

SNA*	SA	-35.009	139.015	8
STG	SA	-35.035	138.571	9
WAG	NSW	-35.128	147.362	9
WAL	NSW	-35.858	146.952	4
WIL	SA	-34.580	138.741	4

*Populations used in the Q_{st}-F_{st} comparison

Table 4 - S3. List of climate variables included in the study and their abbreviations.

Abbreviation	**Description**
var1	Annual mean temperature
var2	Mean diurnal range
var3	Isothermality
var4	Temperature seasonality (standard deviation*100)

var5	Max temperature of the warmest month
var6	Min temperature of the coldest month
var7	Temperature annual range (var5-var6)
var8	Mean temperature of the wettest quarter
var9	Mean temperature of the driest quarter
var10	Mean temperature of the warmest quarter
var11	Mean temperature of the coldest quarter
var12	Annual precipitation
var13	Precipitation of the wettest month
var14	Precipitation of the driest month
var15	Precipitation seasonality (coefficient of variation)

var16	Precipitation of the wettest quarter
var17	Precipitation of the driest quarter
var18	Precipitation of the warmest quarter
var19	Precipitation of the coldest quarter

Table 4 - S4. Correlation (Pearson's) of measured response traits. Highly correlated traits (Pearson's correlation coefficient >0.70) are given in bold font.

	Root length	**Shoot length**	Root to shoot ratio	**Root biomass**	**Shoot biomass**	Total biomass	Proportion of germinated seeds
Root length		**0.71**	**0.76**	**0.72**	0.44	0.58	0.55
Shoot length	**0.71**		0.16	0.64	**0.74**	**0.77**	0.4
Root to shoot ratio	**0.76**	0.16		0.54	0.00	0.18	0.47
Root biomass	**0.72**	**0.77**	0.54		0.6	0.8	0.46
Shoot biomass	0.44	**0.74**	0.00	0.6		**0.94**	0.26
Total biomass	0.58	0.64	0.18	**0.8**	**0.94**		0.36
Proportion of germinated seeds	0.55	0.4	0.47	0.46	0.26	0.36	

Table 4 - S5. Summary of the variation of traits measured in the study.

Trait	Control treatment mean (95% CI)	Temperature treatment mean (95%CI)
Root length	4.92 [4.01, 4.96]	5.57 [4.96, 6.51]
Shoot length	3.98 [3.83, 4.14]	5.32 [5.06, 5.51]
Root to shoot ratio	1.32 [1.13, 1.51]	1.15 [0.97 1.33]
Root biomass	0.83 [0.73, 0.86]	1.32 [1.17, 1.47]
Shoot biomass	0.67 [0.63, 0.70]	0.86 [0.82, 0.96]
Total biomass	1.25 [0.93, 1.29]	1.63 [1.02, 1.76]
Proportion of germinated seeds	0.62 [0.57, 0.68]	0.68 [0.62, 0.74]

Table 4 - S6. Summary of model selection parameters for the linear models built to explain the relationship between trait root length and the first three principal components representing climate.

Model	**df**	**LogLik**	**AICc**	**Delta**	**Weight**
Root length ~ Treatment + PC3	4	-181.36851	371.207614	0	0.2793311
Root length ~ Treatment + PC1 + PC3	5	-180.57461	371.863504	0.65588985	0.20123088

Root length ~ Treatment	3	-183.29124	372.861548	1.65393408	0.12217211
Root length ~ Treatment + PC1	4	-182.39705	373.264694	2.05708053	0.09986882
Root length ~ Treatment + PC2 + PC3	5	-181.36526	373.444808	2.23719413	0.09126805
Root length ~ Treatment + PC1 + PC3	6	-180.55939	374.130821	2.92320754	0.06476686
Root length ~ Treatment + PC1 + PC2 + PC3	4	-183.24458	374.959751	3.75213761	0.04279105
Root length ~ Treatment + PC2	5	-182.31761	375.349515	4.14190105	0.03521413
Root length ~ Treatment + PC1 + PC2	3	-185.18921	376.657483	5.44986906	0.01831029
Root length ~ PC3	4	-184.46045	377.391493	6.18387963	0.01268548
Root length ~ PC1 + PC3	2	-186.95846	378.054848	6.8472339	0.00910461
Root length ~ 1	3	-186.1349	378.548864	7.34125029	0.00711192

Root length ~ PC1	4	-185.18622	378.843027	7.63541357	0.00613918
Root length ~ PC2 + PC3	5	-184.44649	379.607265	8.39965136	0.00418946
Root length ~ PC1 + PC2 +PC3	3	-186.91545	380.109978	8.90236415	0.00325833
Root length ~ PC1 + PC2	4	-186.0618	380.594179	9.38656569	0.00255772

Table 4 - S7. Summary of model selection parameters for the linear models built to explain the relationship between the trait shoot length and the first three principal components representing climate.

Model	**df**	**LogLik**	**AICc**	**Delta**	**Weight**
Shoot length ~ Treatment + PC1	4	-96.758329	201.987246	0	0.33978381
Shoot length ~ Treatment + PC1 + PC2	5	-96.270682	203.255649	1.26840316	0.18020733
Shoot length ~ Treatment + PC1 + PC3	5	-96.323917	203.362119	1.37487323	0.17086486

Shoot length ~ Treatment + PC1 + PC2 + PC3	6	-95.702605	204.417259	2.43001279	0.10081662
Shoot length ~ Treatment	3	-99.29132	204.86171	2.87446358	0.08072738
Shoot length ~ Treatment + PC2	4	-98.679317	205.829222	3.84197624	0.04976547
Shoot length ~ Treatment + PC3	4	-98.765802	206.002193	4.01494651	0.04564236
Shoot length ~ Treatment + PC2 + PC3	5	-97.993067	206.700419	4.71317289	0.03219217
Shoot length ~ PC1	3	-125.50953	257.298123	55.3108764	3.32×10^{-13}
Shoot length ~ 1	2	-126.86435	257.866629	55.8793824	2.50×10^{-13}
Shoot length ~ PC1 + PC2	4	-125.25277	258.976135	56.9888888	1.43×10^{-13}
Shoot length ~ PC1 + PC3	4	-125.28074	259.032065	57.0448187	1.39×10^{-13}

Shoot length ~ PC2	3	-126.53376	259.346585	57.3593393	1.19×10^{-13}
Shoot length ~ PC3	3	-126.58035	259.439769	57.4525232	1.14×10^{-13}
Shoot length ~ PC1 + PC2 +PC3	5	-124.95533	260.624956	58.6377095	6.28×10^{-14}
Shoot length ~ PC2 + PC3	4	-126.16552	260.801629	58.8143825	5.75×10^{-14}

Table 4 - S8. Summary of model selection parameters for the linear models built to explain the relationship between the trait root to shoot ratio and the first three principal components representing climate.

Model	**df**	**LogLik**	**AICc**	**Delta**	**Weight**
Root to shoot ratio ~ PC3	3	-83.099231	172.477531	0	0.21989981
Root to shoot ratio ~ Treatment + PC3	4	-82.233388	172.937363	0.45983204	0.17473246

Root to shoot ratio ~ PC2	4	-82.709992	173.890573	1.41304149	0.10848927
Root to shoot ratio ~ PC1 + PC3	4	-82.91039	174.291367	1.81383579	0.08878823
Root to shoot ratio ~ Treatment + PC2 + PC3	5	-81.836554	174.387393	1.90986198	0.08462596
Root to shoot ratio ~ Treatment + PC1 + PC3	5	-82.04087	174.796025	2.31849378	0.06898749
Root to shoot ratio ~ PC1 + PC2 + PC3	5	-82.47605	175.666386	3.18885465	0.04464511
Root to shoot ratio ~ 1	2	-85.830459	175.798848	3.32131684	0.041784
Root to shoot ratio ~ Treatment + PC1 + PC2 + PC3	6	-81.598014	176.208077	3.73054558	0.03405236
Root to shoot ratio ~ Treatment	3	-85.016068	176.311206	3.83367458	0.03234097

Root to shoot ratio ~ PC2	3	-85.201599	176.682267	4.20473556	0.02686446
Root to shoot ratio ~ Treatment + PC2	4	-84.375642	177.221871	4.74433993	0.02051184
Root to shoot ratio ~ PC1	3	-85.571507	177.422084	4.94455282	0.0185579
Root to shoot ratio ~ Treatment + PC1	4	-84.752374	177.975336	5.49780442	0.01407317
Root to shoot ratio ~ PC1 + PC2	4	-84.880434	178.231456	5.75392493	0.01238158
Root to shoot ratio ~ Treatment + PC1 + PC2	5	-84.048506	178.811298	6.3337664	0.00926542

Table 4 - S9. Summary of model selection parameters for the linear models built to explain the relationship between the trait root biomass and the first three principal components representing climate.

Model	**df**	**LogLik**	**AICc**	**Delta**	**Weight**

Root biomass ~ Treatment	3	-75.521143	157.321355	0	0.19416492
Root biomass ~ Treatment + PC3	4	-74.484844	157.440276	0.11892113	0.1829563
Root biomass ~ Treatment + PC1	4	-74.756531	157.98365	0.66229486	0.13942969
Root biomass ~ Treatment + PC2	5	-73.78768	158.289645	0.96828987	0.11964906
Root biomass ~ Treatment + PC2 + PC3	5	-73.79585	158.305985	0.98463051	0.11867547
Root biomass ~ Treatment + PC1 + PC3	4	-75.031067	158.532723	1.21136832	0.10595598
Root biomass ~ Treatment + PC2	6	-73.184517	159.381082	2.05972758	0.06932767
Root biomass ~ Treatment + PC1 + PC2 + PC3	5	-74.335548	159.385382	2.06402738	0.06917878

Root biomass ~ Treatment + PC1 + PC2	2	-83.777113	171.692157	14.370802	0.00014709
Root biomass ~ 1	3	-82.916192	172.111453	14.7900985	0.00011927
Root biomass ~ PC3	3	-83.141575	172.56222	15.2408648	9.52×10^{-05}
Root biomass ~ PC1	3	-83.369557	173.018185	15.6968296	7.58×10^{-05}
Root biomass ~ PC2	4	-82.338911	173.14841	15.8270551	7.10×10^{-05}
Root biomass ~ PC2 + PC3	4	-82.345667	173.161923	15.840568	7.05×10^{-05}
Root biomass ~ PC1 + PC3	4	-82.79244	174.055467	16.7341125	4.51×10^{-05}
Root biomass ~ PC1 + PC2 + PC3	5	-81.840714	174.395714	17.074359	3.81×10^{-05}

Table 4 - S10. Summary of model selection parameters for the linear models built to explain the relationship between the trait shoot biomass ratio and the first three principal components representing climate.

Model	df	LogLik	AICc	Delta	Weight
Shoot biomass ~ Treatment + PC2 + PC3	5	62.8027839	-114.89128	0	0.50316791
Shoot biomass ~ Treatment + PC1 + PC2 + PC3	6	63.1177523	-113.22346	1.6678257	0.21854918
Shoot biomass ~ Treatment + PC1 + PC2	4	60.6634382	-112.85629	2.03499385	0.18189453
Shoot biomass ~ Treatment + PC3	5	61.0695713	-111.42486	3.46642523	0.08891771
Shoot biomass ~ Treatment	4	56.3818518	-104.29312	10.5981667	0.00251391
Shoot biomass ~ Treatment + PC1 + PC3	3	55.1886112	-104.09815	10.7931295	0.00228042
Shoot biomass ~ Treatment + PC2 + PC3	5	56.8735807	-103.03288	11.8584063	0.00133873

Shoot biomass ~ Treatment + PC1	4	55.7508994	-103.03121	11.8600715	0.00133762
Shoot biomass ~ PC2 + PC3	4	40.3567804	-72.242973	42.6483094	2.76×10^{-10}
Shoot biomass ~ PC2	3	39.0455577	-71.812046	43.0792365	2.22×10^{-10}
Shoot biomass ~ PC1 + PC2 + PC3	5	40.5477848	-70.381284	44.5099982	1.09×10^{-10}
Shoot biomass ~ PC1 + PC2	4	39.2963335	-70.122079	44.7692033	9.55×10^{-11}
Shoot biomass ~ 1	2	35.5813855	-67.02484	47.8664422	2.03×10^{-11}
Shoot biomass ~ PC3	3	36.3495516	-66.420034	48.4712485	1.50×10^{-11}
Shoot biomass ~ PC1	3	35.9442707	-65.609472	49.2818104	1.00×10^{-11}

Shoot biomass ~ PC1 + PC3	4	36.6639933	-64.857398	50.0338838	6.87×10^{-12}

Table 4 - S11. Summary of model selection parameters for the linear models built to explain the relationship between the trait total biomass and the first three principal components representing climate.

Model	**df**	**LogLik**	**AICc**	**Delta**	**Weight**
Total biomass ~ Treatment + PC2 + PC3	5	-27.349216	65.412717	0	0.38524557
Total biomass ~ Treatment + PC1 + PC2 + PC3	6	-26.425964	65.863975	0.45125815	0.30743138
Total biomass ~ Treatment + PC2	4	-29.608677	67.687941	2.27522456	0.1235034
Total biomass ~ Treatment + PC1 + PC2	5	-28.550405	67.815096	2.40237888	0.1158958
Total biomass ~ Treatment + PC1 + PC3	5	-30.316497	71.347280	5.93456306	0.01981817
Total biomass ~ Treatment + PC3	4	-31.445267	71.361121	5.94840442	0.01968149

Total biomass ~ Treatment + PC1	4	-31.70532	71.881228	6.46851118	0.01517464
Total biomass ~ Treatment	3	-32.936742	72.152554	6.73983669	0.01324954
Total biomass ~ PC2 + PC3	4	-47.708252	103.88709	38.4743754	1.70×10^{-09}
Total biomass ~ PC2	3	-49.158531	104.59613	39.1834149	1.19×10^{-09}
Total biomass ~ PC1+ PC2 + PC3	5	-47.123182	104.96064	39.5479319	9.95×10^{-10}
Total biomass ~ PC1 + PC2	4	-48.476013	105.42261	40.0098965	7.90×10^{-10}
Total biomass ~ 1	2	-51.34175	106.82143	41.4087136	3.93×10^{-10}
Total biomass ~ PC3	3	-50.356473	106.99201	41.5792994	3.61×10^{-10}

Total biomass ~ PC1	3	-50.527467	107.33400	41.9212873	3.04×10^{-10}
Total biomass ~ PC1 + PC3	4	-49.618203	107.70699	42.2942767	2.52×10^{-10}

Table 4 - S12. Summary of model selection parameters for the linear models built to explain the relationship between the trait proportion of germinated seeds and the first three principal components representing climate.

Model	**df**	**LogLik**	**AICc**	**Delta**	**Weight**
Proportion of germination ~ PC1	3	16.8341001	-27.38913	0	0.22209765
Proportion of germination ~ PC1 + PC2	4	17.7746116	-27.078635	0.31049534	0.19016069
Proportion of germination ~ PC1	4	17.0944132	-25.718238	1.67089216	0.09631953
Proportion of germination ~ PC1 + PC3	4	16.8914276	-25.312267	2.07686342	0.07862466
Proportion of germination ~ Treatment + PC1	5	17.934731	-25.155176	2.23395419	0.07268536

Proportion of germination ~ PC1+ PC2 + PC3	5	17.8331507	-24.952016	2.43711466	0.06566459
Proportion of germination ~ Treatment + PC1 + PC2	3	15.4976257	-24.716182	2.67294878	0.0583607
Proportion of germination ~ PC2	2	14.4050716	-24.672212	2.71691819	0.05709166
Proportion of germination ~ 1	5	17.1520736	-23.589861	3.799269	0.03323098
Proportion of germination ~ Treatment + PC1 + PC3	6	17.9934788	-22.974909	4.41422091	0.02443476
Proportion of germination ~ Treatment + PC1 + PC2 + PC3	3	14.5754794	-22.871889	4.51724127	0.02320799
Proportion of germination ~ PC3	4	15.5855939	-22.7006	4.6885308	0.02130308
Proportion of germination ~ PC2 + PC3	3	14.4593849	-22.6397	4.74943034	0.02066418

Proportion of germination ~ Treatment	4	15.5532746	-22.635961	4.75316939	0.02062559
Proportion of germination ~ Treatment + PC2	4	14.6299989	-20.78941	6.59972077	0.00819281
Proportion of germination ~ Treatment + PC2 + PC3	5	15.6413518	-20.568418	6.82071251	0.00733576

Table 4 - S13. Summary of the calculated standard measures of genetic diversity per population averaged across loci.

Population	**H_E**	**H_o**	**A_R**	**A_R 95% CI**
ASH	0.08	0.12	1.07	[0.62, 1.13]
BARK	0.05	0.07	0.95	[0.47, 1.02]
BOR	0.09	0.12	1.09	[0.74, 1.22]
BUR	0.05	0.07	1.09	[0.91, 1.20]
CALL	0.07	0.13	1.12	[0.84, 1.19]

COB	0.04	0.06	0.96	[0.46, 1.08]
CRC	0.12	0.09	0.96	[0.42, 1.06]
GOR	0.05	0.07	1.00	[0.46, 1.12]
GOU	0.09	0.07	0.96	[0.37, 1.05]
HOV	0.04	0.07	1.06	[0.85, 1.14]
IND	0.06	0.08	1.12	[0.93, 1.18]
KEI	0.06	0.08	1.06	[0.67, 1.20]
LAV	0.07	0.09	1.04	[0.58, 1.16]
MAT	0.05	0.07	1.08	[0.95, 1.17]
MOA	0.06	0.08	1.08	[0.81, 1.17]
MUR1	0.07	0.14	1.11	[0.74, 1.22]
OSM	0.07	0.08	0.94	[0.47, 1.07]

RAI	0.07	0.12	1.07	[0.68, 1.17]
ROD	0.04	0.06	1.04	[0.93, 1.07]
RUG	0.07	0.11	1.07	[0.73, 1.13]
SNA	0.05	0.07	1.02	[0.66, 1.11]
STG	0.06	0.10	1.05	[0.70, 1.16]
WAG	0.07	0.08	1.12	[0.90, 1.20]
WAL	0.07	0.10	1.11	[0.98, 1.16]
WIL	0.09	0.08	0.95	[0.65, 1.06]

A B

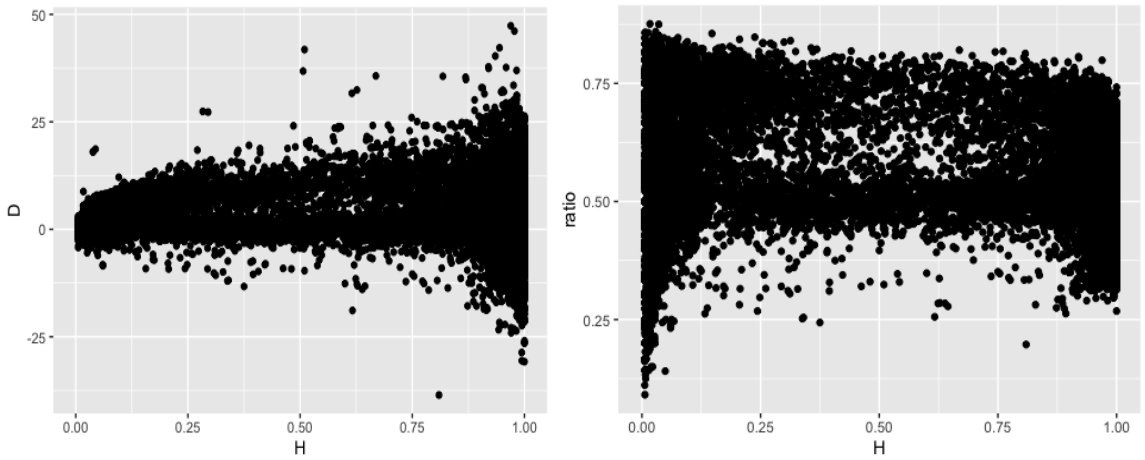


Figure 4 - S1. Plots from HDPlot package. (A). Plot of the proportion of heterozygotes (H) against read ratio deviation (D). (B). plot of the proportion of heterozygotes against the read depth ratio of the two alleles of the heterozygous individuals.

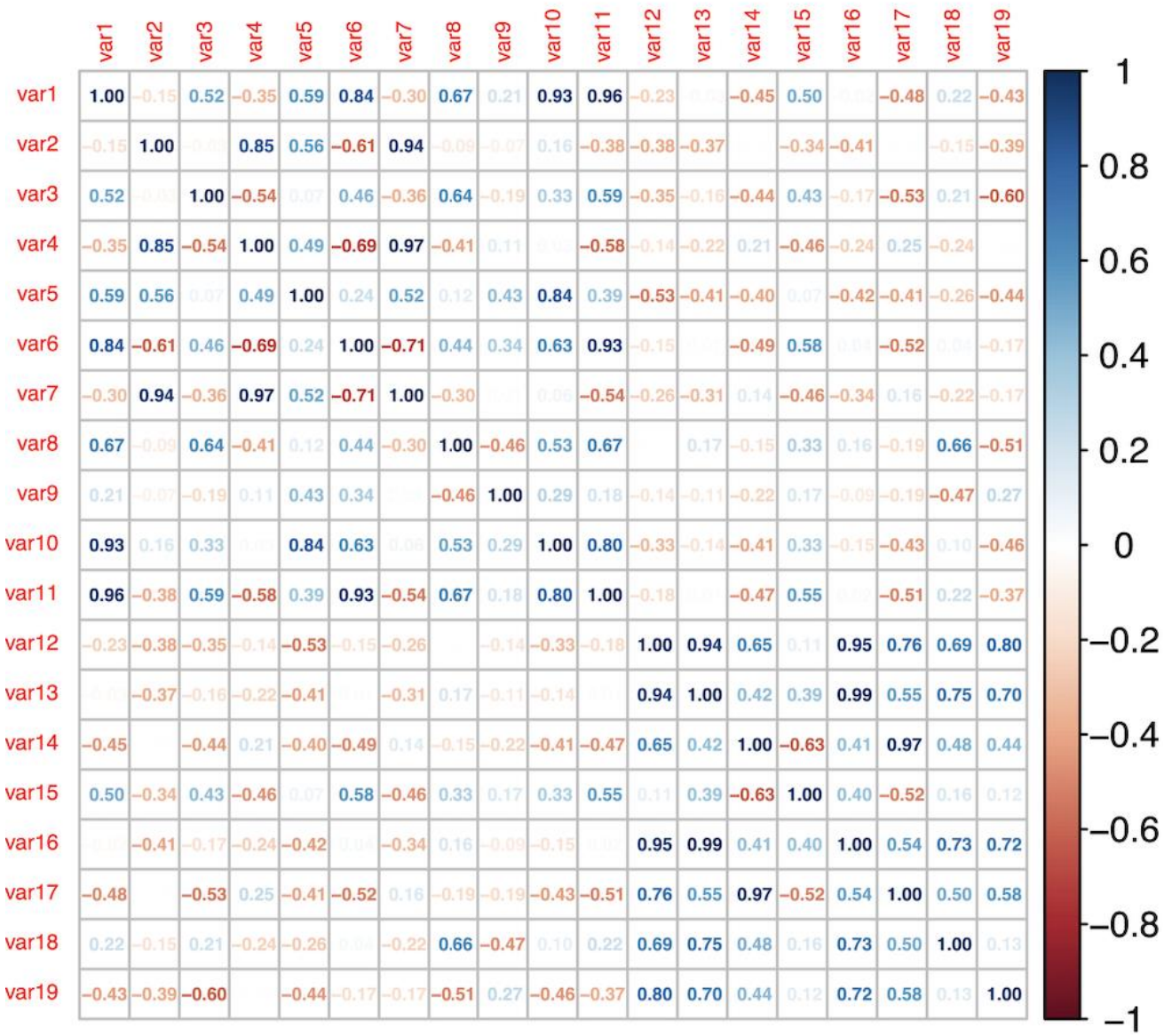


Figure 4 - S2. Corrplot depicting the correlation matrix of the 19 bioclimatic variables used in the phenotypic trait analysis.

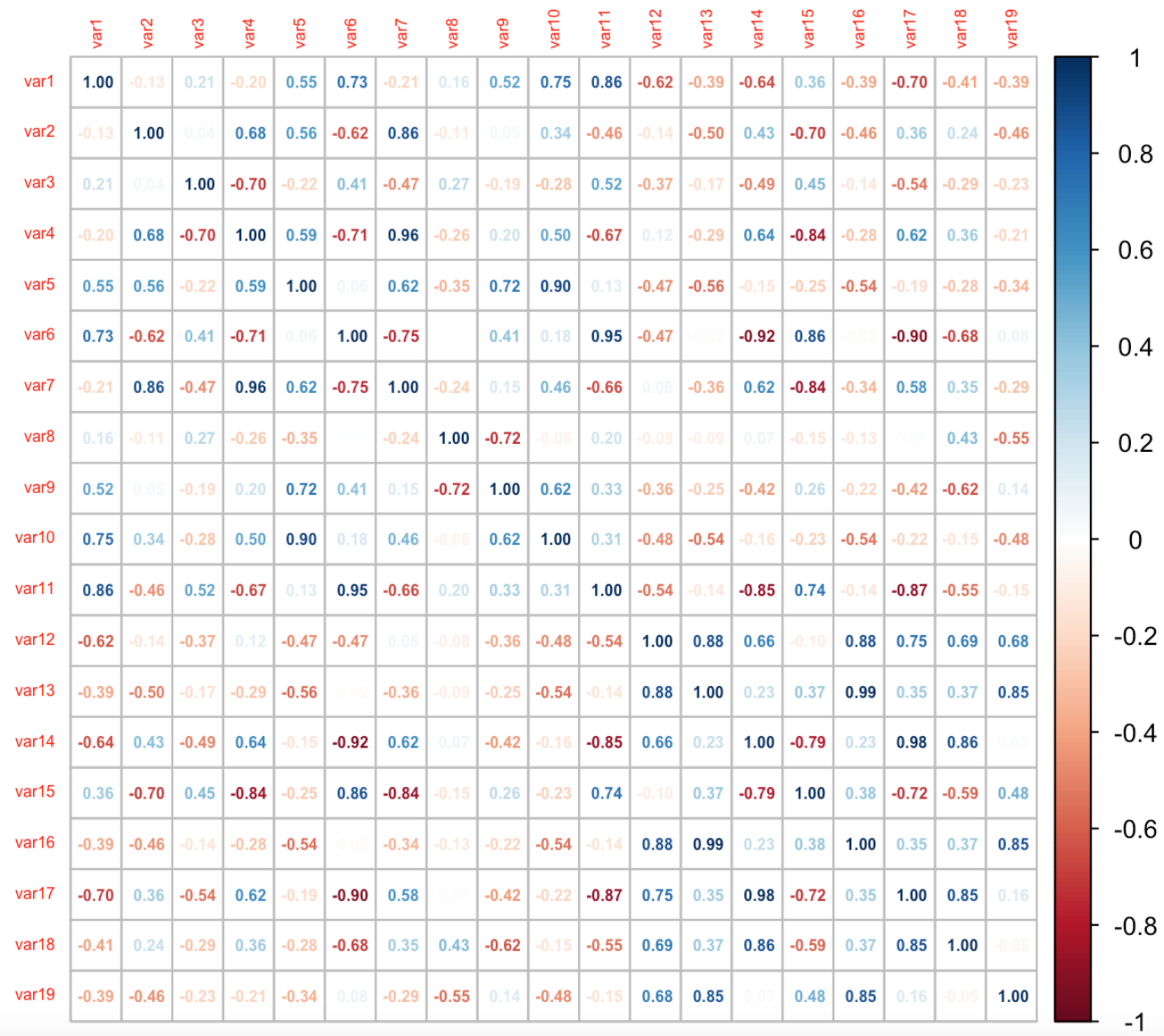


Figure 4 - S3. Corrplot depicting the correlation matrix of the 19 bioclimatic variables used in the isolation by environment (IBE analysis).

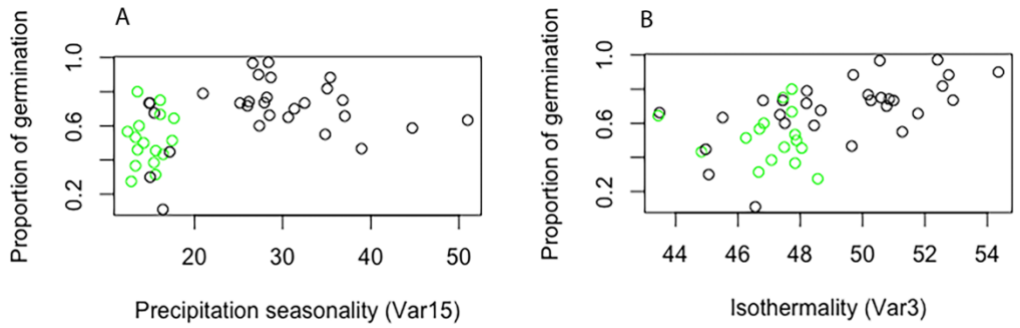


Figure 4 - S4. Biplots of variation of proportion of germination in control treatment with precipitation seasonality and isothermality. (A) with precipitation seasonality (B) with isothermality. Trait values in populations found in the red region in Figure 4 - 4(A) are highlighted in green.

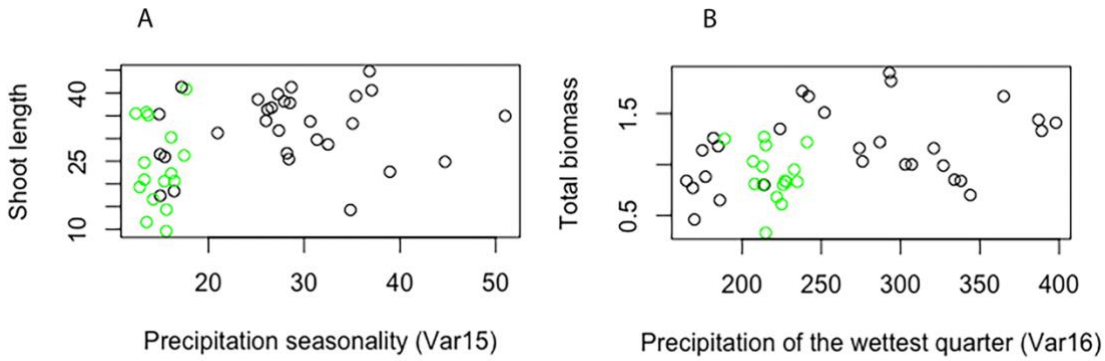
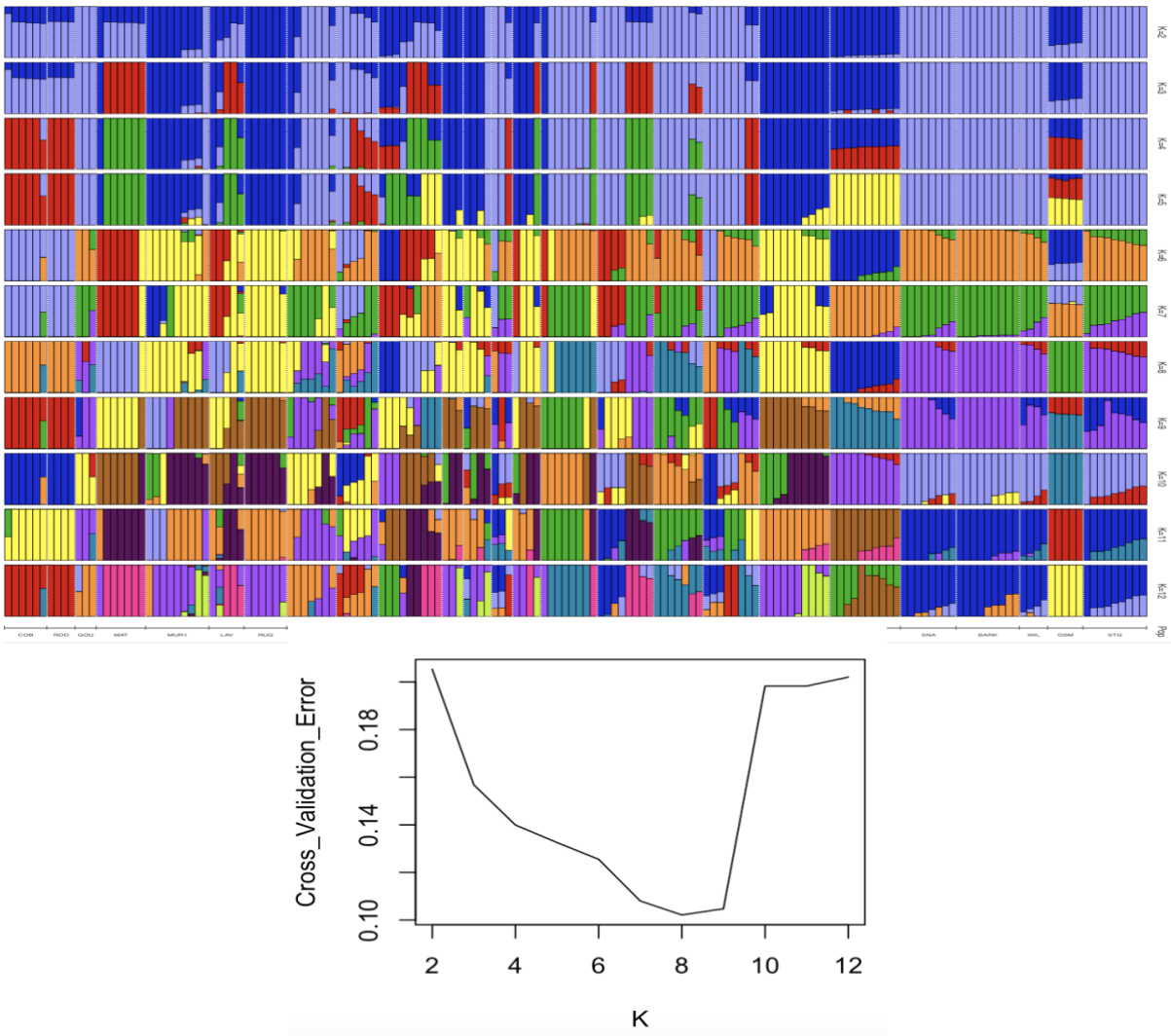


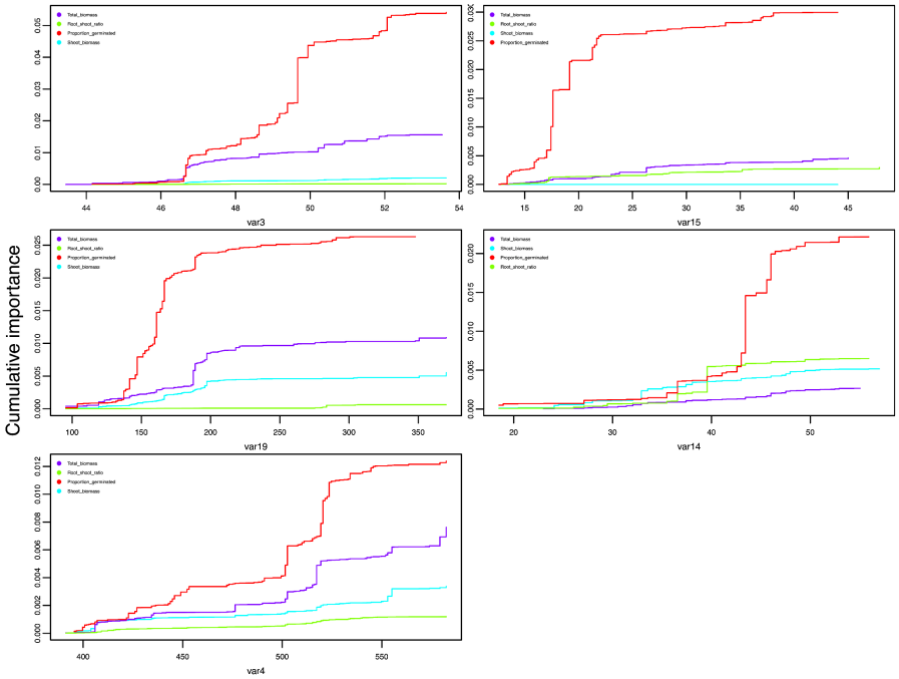
Figure 4 - S5. Biplots of variation of shoot length and total biomass in temperature treatment with climate variables. (A) Variation of shoot length with precipitation seasonality (B) variation of total biomass with precipitation of the wettest quarter. Trait values in populations found in the green region at the centre of the range in Figure 4 - 4(C) are highlighted in green.



B

Figure 4 - S6. (A) A distruct plot showing the results of Admixture run of the for filtered SNP set for k=2-12. (B) Plot of the distribution of the cross-validation score for each k value.

A



B

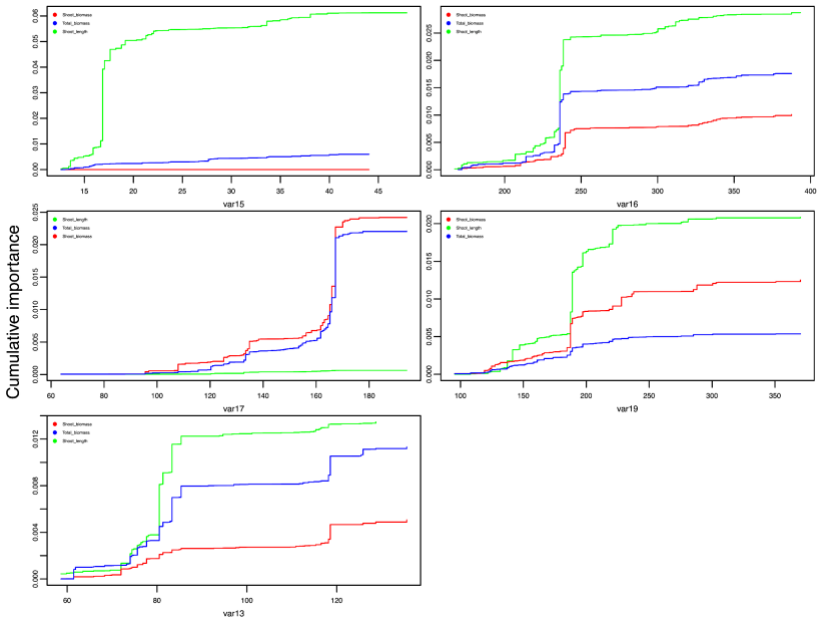


Figure 4 - S7. Cumulative distribution importance plots from the gradient forest models. (A) Cumulative importance distribution of the top traits used in the gradient forest run with traits measured under control treatment with relation to the top five important climate variables

predicting the phenotypic variation. Colours representing traits are, purple-total biomass, green-root to shoot ratio, red-proportion of germinated seeds and blue-shoot biomass. (B) Cumulative importance distribution of the top traits used in the gradient forest run with traits measured under temperature treatment with relation to the top five important climate variables predicting the phenotypic variation. Colours representing traits are, red-shoot biomass, blue-total biomass, and green shoot length.

Chapter 5 - General Discussion

This thesis includes one of the few studies investigating the evolutionary and genetic potential of ecologically important Australian grassland species and assessing their suitability to be used in degraded grassland restoration. Two native Australian grasses, *Bothriochloa decipiens* and *Bothriochloa macra*, were the focal species used in my thesis. The main aims of this thesis were to: 1) investigate the evolutionary significance of whole genome duplication/allopolyploidy; 2) develop genomic resources for future Australian grass genomic studies; and 3) investigate the population structure and signals of climate adaptation of native Australian grasses. To achieve these aims I sampled individuals of *B. decipiens* and *B. macra* populations from across Australia. I conducted population genomic analysis and GEA analysis using *B. decipiens* genomic data and a population genomic analysis was done using data from *B. macra*. I collected and analysed phenotypic data from *B. macra* seedlings in a common garden experiment. Finally, I generated a chromosomal level genome assembly and annotation for *B. decipiens* followed by an extensive comparative analysis. This research will enhance the availability of genomic resources and provide insights into the genetic diversity and evolutionary potential of ecologically important grass species. Most importantly, findings from this research shed light on the hypothesis that allopolyploidy facilitates adaptation to new environments, thereby enabling successful expansion and establishment of C4 grasses.

Evolutionary potential of the two grass species and restoration implications

Genetic structuring and variation within species is identified as an important driver of grassland functioning and ecosystem services (Greenville et al., 2017; van der Meer & Jacquemyn, 2015). Identifying genetic patterns is crucial for conservation and restoration (Zhai et al., 2015) because informing restoration management activities with knowledge on genetic potential can guide restoration efforts towards sites of unique genetic composition or adaptive importance (Ikeda et al., 2017; McCartney-Melstad & Shaffer, 2015; Reynolds et al., 2015). Further, characterising population structure is the key in identifying the relative contribution of different evolutionary processes (geneflow, drift and selection) across populations (Hohenlohe et al., 2021). Understanding the amount of gene flow among populations through population structure is important particularly during restoration of species in isolated or fragmented landscapes (Crooks & Sanjayan, 2006; Walters & Schwartz, 2021). When human mediated transfer of propagules to assist dispersal and increase connectivity among populations by increasing the gene flow occurs, there is always the risk of maladaptation of the new genotypes and also the new genotypes might interbreed with the local population and produce offspring with lower fitness (outbreeding depression) (Hufford & Mazer, 2003). Therefore, investigating the population structure of restoration species is important to ensure that potential seed transfer occurs between populations with similar genetic structure and variability (Broadhurst et al., 2008).

Both my study species, *Bothriochloa decipiens* and *Bothriochloa macra* populations showed very high population differentiation and signals of isolation by distance (Chapter 3 and 4). The extent of differentiation among populations in different species can vary and depend on many evolutionary processes such as gene flow, genetic drift, mutations, and selection (Slatkin, 1987). Further, the importance of these forces may be impacted by strength of selection, population size, barriers for dispersal and plant life history traits such as the mating system (Loveless & Hamrick, 1984). As *B. decipiens* is known to be frequently cleistogamous in nature (Connor, 1979) and *B. macra* has both chasmogamous and cleistogamous flowers in varying proportions (Yu et al., 2000), and as cleistogamy leads to self-fertilisation, this could be one factor contributing to population genetic clustering among the populations in *B. decipiens* in particular. Cleistogamous mating system and high self-fertilisation has accounted for the presence of population structuring in other grass species (Lowry et al., 2013; Parkhurst et al., 2011). In

addition, apomixis and other forms of asexual reproduction are common in grasses and can contribute to population genetic structure (Brugnoli et al., 2014; Thammina et al., 2018). Strong genetic clustering is reported in another Australian native grass species *Themeda triandra* (Ahrens et al., 2020*)* but the F_{st} reported in this study is lower than the differentiation reported in this study system. However, my analysis of selfing rates in *B. decipiens* did not reveal uniformly high selfing rates and clonal genotypes were also not detected in the samples from either species. Still, it is likely that the mating system of both species contributed to the population genetic structuring I observed.

Genetic diversity and structure of plant populations are strongly influenced by gene flow occurring through pollen and seed dispersal (Wright, 1984). Pollen mediated gene movement affects within and among population genetic diversity in plants, whereas seed dispersal plays a role in colonisation, reestablishment, and local migration (Barluenga et al., 2011). The two grass species in this study are wind pollinated. Size, density and distribution of both donor and acceptor populations are important for gene flow in wind pollinated plants (Rognli et al., 2000). Neighbouring plants received most of the pollen released in both wind and insect pollinated plants (Gleaves, 1973; Handel, 1983). Studies of gene flow dynamics of grass and other wind pollinated species reveal that although pollen can be dispersed over a large distance via wind pollination, gene flow is often much more restricted. Most pollen deposition occurs close to the donor plant and plant density greatly influences gene flow (Rognli et al., 2000). Even a minute change in plant density can affect pollination success, further many factors such as plant height, flower position and characteristics of stamen and stigma affect the success of pollen dispersal and capture (Friedman & Barrett, 2009). The primary mode of seed dispersal in the two grass species I studied is by adhesion via animals and this could mean more restricted and concentrated seed dispersal within fragmented geographic locations in the recent past (Bittencourt & Sebbenn, 2007; Ghazoul, 2005), although dispersal could have happened over large distances with the help of animals such as kangaroos in the past. Indeed, perhaps rare long distance seed dispersal, caused by humans or other animals, explains the mixed populations in the Admixture analysis. Combined, the effects of these limitations in pollen and seed dispersal could be another factor affecting the observed genetic structure in the two study species.

Heterogeneous environments help in maintaining ecologically important genetic variation among populations (McKay et al., 2001). Many studies on grasses have discovered genetic differentiation among populations associated with both geography and environmental

heterogeneity (Arnesen et al., 2017; Gould et al., 2018; Guo et al., 2021; Liu et al., 2016; Lowry et al., 2013; Marques et al., 2017; Oyundelger et al., 2020; Šurinová et al., 2019). The two grasses in this study are also widely distributed across eastern Australia and face different environments and climates, this could also contribute to the presence of population genetic structure. However, separating the environmental and geographic effects on population structure is challenging as these factors are frequently intercorrelated.

Populations spread across environment gradients have the potential to develop local adaptation to the environment (Conner et al., 2004) and local adaptation is quite common (Hereford, 2009; Leimu & Fischer, 2008). Many studies have demonstrated that grasses have a great capacity to adapt to local environments (Arnesen et al., 2017; Gray et al., 2014; Hamann et al., 2016; Liu et al., 2016; Meyer et al., 2016; Rudmann-Maurer et al., 2007). Similarly, both *B. decipiens* and *B. macra* provided evidence consistent with patterns of local adaptation. Genomic signatures of climate adaptation in *Bothriochloa decipiens,* especially for loci related to drought stress response, were detected via landscape genomic analysis in chapter 3. In addition, *B. macra* seedlings from high rainfall and low temperature areas yielded higher total biomass values when compared with seedlings from low rainfall and high temperature areas, consistent with the hypothesis that populations from more arid regions have evolved a 'drought tolerance' mechanism to survive drought by decreasing resource investment in overall growth and total biomass allocation (Chapter 4). However, non-adaptive explanations are also possible given the significant population structure observed in this species. As the presence of local adaptation to climate factors is an important determinant when selecting candidates for practising assisted gene flow in restoration (Aitken & Whitlock, 2013), these signals of local adaptation discovered in the two grass species may suggest their suitability to be used in climate change resilient grassland restoration. Moderate population structuring and signals of IBD suggest limited dispersal in both the grass species and may point to the importance of mixing propagules through assisted gene flow as climate adaptation through natural gene flow may be limited.

Polyploidy and local adaptation

Numerous studies associate polyploidization with adaptation to abiotic stress such as extremely dry and cold environments (Ehrendorfer, 1979; Folk et al., 2020; Levin, 1983; Lourkisti et al., 2020). Some studies argue that allopolyploidy gave grasses the ability to adapt to new

environments, thereby enabling successful expansion and establishment (Godfree et al., 2017; Linder & Barker, 2014). Also, in *Themeda triandra*, signals for ploidy based adaptive response were discovered as polyploids are dominant in hot and dry regions (Ahrens et al., 2020). In the genome of the diploidized allopolyploid *B. decipiens* I discovered biased retention of duplicated genes related to stress response in chapter 2. Further in chapter 3, I discovered that the proportion of climate adaptation candidate genes present as duplicates was higher than the proportion present as single copy. Integrated findings from chapter 2 and 3 support the argument that whole genome duplication through hybridization can facilitate adaptation to new environments.

Biassed retention of genes with regulatory and developmental functions after a WGD is common knowledge (Blanc & Wolfe, 2004; Maere et al., 2005). In some cases these genes are retained as duplicates not necessarily due to any selective advantage but because they are important to maintain gene stoichiometry balance (Birchler & Veitia, 2012; Freeling, 2009). However, with sufficient time these genes may evolve to perform new functions facilitating adaptation to environmental variation. In a recent study stimulating biological evolution, artificial non-polyploids performed better than polyploids in more stable conditions whereas artificial polyploids performed better in unstable conditions (Yao et al., 2019). Random mutations cause changes in genomes and gene regulatory networks, these changes can be amplified in genomes that have undergone WGDs. In stable environments these random mutations in polyploid genomes should cause maladaptation and can be detrimental, but under unstable stressful conditions these complexities may provide the substantial variation necessary for adaptation and survival (Carretero-Paulet & Van de Peer, 2020; Yao et al., 2019). As neofunctionalization and specialisation can be possible fates of genes retained as duplicates, retention of such genes may promote adaptive evolution to abiotic stresses (Cheng et al., 2018; Defoort et al., 2019; Jiao et al., 2018; Panchy et al., 2016; Ren et al., 2018; Soltis & Soltis, 2016).

Subgenome dominance (Flagel & Wendel, 2010) after a whole genome duplication where one subgenome retains more genes and has higher gene expression is common among allopolyploids (Chen et al., 2020; Edger et al., 2019; Session et al., 2016). Stronger signals of climate adaptation might be expected to be seen in the dominant subgenome, due to its high expression and greater gene retention (Bird et al., 2018). Different observations relating the contributions of the subgenomes to adaptation are presented in literature (Lovell et al., 2021; Wang et al., 2022; Zhang et al., 2015). For example, in switchgrass the dominant subgenome showed reduced

heritability in SNPs for biomass and survival (Lovell et al., 2021). The authors suggested that reduced purifying selection acting on the non-dominant subgenome may allow for the accumulation of new variation which provides the fuel for subsequent adaptation. In Zoysiagrass, no significant biassed fractionation was found between the two subgenomes, but there was a homologous expression bias in one subgenome in pathways involved in abiotic stress response which may reflect adaptation to adverse environments (Wang et al., 2022). Similarly, in some allotetraploids it has been found that different subgenomes contributed to different characteristics (Zhang et al., 2015). In chapter 2 I identified the dominant subgenome of *B. decipiens* as the subgenome with greater gene retention. In chapter 3, I tested if there is a bias in the presence of climate adaptation candidate genes between the two subgenomes but failed to detect any association.

Future directions

In this thesis I present crucial information on genomics of two ecologically important native grass species, and it is one of the few genomic studies of an Australian understory species. Although I present many exciting insights on the evolutionary and genomic potential of these grasses, many improvements and incorporation of different research methodologies could help further our understanding.

In chapter 3 I identified climate adaptation in *Bothriochloa decipiens* for loci related to drought stress through landscape genomic analysis. These findings were based on genomic data generated using just eight populations and to arrive at more robust conclusions the number of sampling populations should be increased. Further, I discovered that this species is distributed across an aridity gradient and my sampling locations do not represent samples from more arid regions in the distribution therefore further sampling to include populations from more arid regions should be done. I could not sample more broadly or conduct common garden trials to combine with genomic data due to time constraints amplified by the bushfires that occurred during planned seed collection trips and the recent Covid-19 pandemic which had significant impacts on travel and research in Melbourne, but for a more robust analysis this chapter could be improved by combining phenotypic and genetic environment association analysis across a broader panel of accessions. In some studies, phenotypic and/or functional trait data is collected during sampling on site of the same accessions used to generate genomic data to conduct

analysis combining phenotype, genotype, and environmental data (Carvalho et al., 2021; Steane et al., 2017). However, a powerful approach to examine genetic divergence in traits, would be to conduct common garden trials to obtain phenotypic data in standard environments and to combine this with genomic analysis (Faske et al., 2021; Fitzpatrick et al., 2021; Mahony et al., 2019; Steane et al., 2014; Jordan et al., 2020). Combining such phenotypic and genomic data to study local adaptation is seen in few studies focussing on non-domesticated grass species (Liu et al., 2016; López et al., 2020; Lowry et al., 2013).

In chapter 3, I used landscape genomics to predict the genetic vulnerability of *B. decipiens* populations to future climate change. Landscape genomics is used to predict how populations will respond to climate change across a species' range and assess the importance of local adaptation in these responses (Exposito-Alonso et al., 2018; Fitzpatrick & Keller, 2015), but adaptation alone will not determine the evolutionary fates of populations because evolution is influenced by many interrelated factors such as mutation, genetic drift, geneflow, and recombination (Reside et al., 2018). Developing means to integrate evolutionary forces in species' response to climate change predictions can assist in conservation and management (Allendorf et al., 2010; Brown et al., 2016; Waldvogel et al., 2020). Consequently, my predictions could be further improved by more explicitly including the effects of evolutionary forces by following a framework as suggested in (Aguirre-Liguori et al., 2021). They describe the use of the FOLDS model, an integrated framework to incorporate gene flow, genetic offset, dispersal, genetic load, and species distribution models when predicting species' response to future change.

Integrating experimental common garden and landscape genomic data with simulations could improve our understanding of the evolution of adaptive traits, population persistence under climate change and even future species distributions (e.g., Bush et al., 2016). Such evolutionarily informed species distribution modelling should be applied to species used in ecological restoration to predict their fate under climate change. Further, such simulations could explicitly incorporate various AGF strategies to predict the usefulness of these approaches (Christmas et al., 2016; Fiedler et al., 2018). Such predictions can then be tested using field experiments. Given the foundational work I have conducted in *Bothriochloa,* including the high-quality reference genome and population genomic analysis, these species are excellent candidates for this approach in the future.

Bothriochloa macra is likely an allotetraploid (Sumadijaya, 2015), and during the SNP calling for my fourth chapter, I aligned the GBS reads from *B. macra* to the genome of the diploidized allopolyploid *B. decipiens*. I could not calculate selfing rates of populations because inclusions of paralogs could strongly bias the results and carrying out downstream landscape genomic analysis was not possible because of the excessive loss of data during filtering for heterozygosity and the possible biases introduced during genotyping that could cause false positive outlier loci. These issues could be resolved if an allopolyploid genome reference for *B. macra* could be assembled. Phasing approaches to construct haplotype resolved genome assemblies by separating haplotype sequences from the subgenomes, alongside those inherited from the maternal and paternal parent (Edge et al., 2017; Koren et al., 2018; Kronenberg et al., 2021). For example, in a recent review on the operating principles and characteristics of freely available phasing tools (Guk et al., 2022), ALLHiC (Zhang et al., 2019) was recognized as a phasing tool for generating haplotype resolved assemblies for complex polyploid species. Therefore, due to the development of high quality long reads and the computational advances in assembly, it is now feasible to construct high quality assemblies for complex allopolyploid genomes like that of *B. macra* and this should greatly advance the population genomic analysis.

Developing a genome assembly for *B. macra* and doing a similar comparative analysis on it as the one in chapter 2 would allow a comparison between the two whole genome duplication events and provide insight into how fractionation and biassed gene retention proceed over time in the group. The genus *Bothriochloa* belongs to the BCD clade, with a complex history of hybridization (De Wet & Harlan, 1966; Harlan & de Wet, 1963). Therefore, the two genomes of these closely related *Bothriochloa* species could be used to look for any patterns of introgression. As introgression can be a source of new genetic variation and result in adaptation (Abbott et al., 2013; Rieseberg et al., 2003; Schmickl et al., 2017), this could further our knowledge on the adaptive potential of this grass genus and its importance for climate resilient restoration.

Extensive common garden experiments that span across a greater time, and allow us to obtain phenotypic data in matured individuals (Galliart et al., 2019; von Wettberg et al., 2014) could be conducted for both species. Range wide provenance trials, accompanied by population genomic analysis, could be designed to comprehensively assess signals of climate adaptation in these species. Further the use of long-term provenance trials would allow the development and evaluation of seed transfer zones (Aitken et al., 2008; Hamann et al., 2011; Nagamitsu et al.,

2014; Rehfeldt et al., 1999). Such trials could provide great insight into the fitness effects and demographic effects of seed transfer and even allow assessment of the predictions of the Gradient Forest analyses (Láruson et al., 2022). Since selfing and the production of clones is possible, ideally, we could develop a genetically diverse panel of lineages from a wide array of climates to use in these trials. This would allow us to reduce sequencing costs and to specifically examine plasticity in response to environmental variation.

Conclusion

In this thesis I present novel findings on the ecological and evolutionary potential of two Australian native grass species. Both species show signals of local adaptation to climate and could be considered for climate change resilient restoration of grassland ecosystems in eastern Australia. Furthermore, my assembly of *B. decipiens'* complex and large genome, allowed me to investigate the role of whole genome duplications/polyploidy in genome evolution and adaptation in plants. My findings in chapter 2 and 3 support the argument that gene duplication occurring via polyploidy helps plants to acquire adaptation to new conditions. The results from this thesis will help to fill the gap related to the availability of knowledge and resources on the genetic potential of ecologically important Australian understory species.

Bibliography

Aarssen, L. W., & Burton, S. M. (1990). Maternal effects at four levels in *Senecio vulgaris* (Asteraceae) grown on a soil nutrient gradient. *American Journal of Botany*, *77*(9), 1231–1240.

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246.

Aguirre-Liguori, J. A., Ramírez-Barahona, S., & Gaut, B. S. (2021). The evolutionary genomics of species' responses to climate change. *Nature Ecology & Evolution*, *5*(10), 1350–1360.

Ahrens, C. W., James, E. A., Miller, A. D., Scott, F., Aitken, N. C., Jones, A. W., Lu-Irving, P., Borevitz, J. O., Cantrill, D. J., & Rymer, P. D. (2020). Spatial, climate and ploidy factors drive genomic diversity and resilience in the widespread grass *Themeda triandra*. *Molecular Ecology*, *29*(20), 3872–3888.

Aitken, S. N., & Bemmels, J. B. (2016). Time to get moving: assisted gene flow of forest trees. *Evolutionary Applications*, *9*(1), 271–290.

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*(1), 367–388.

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*(1), 95–111.

Alexa, A., & Rahnenfuhrer, J. (2006). TopGO: enrichment analysis for gene ontology. *R package version*.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664.

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews. Genetics*, *11*(10), 697–709.

All Tansposase Protein Database. Retrieved November 13, 2020, from http://www.hrt.msu.edu/uploads/535/78637/Tpases020812.gz

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410.

Andrew, R. L., Ostevik, K. L., Ebert, D. P., & Rieseberg, L. H. (2012). Adaptation with gene flow across the landscape in a dune sunflower. *Molecular Ecology*, *21*(9), 2078–2091.

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., & Ogata, H. (2020). KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* , *36*(7), 2251–2252.

Arnesen, S., Coleman, C. E., & Meyer, S. E. (2017). Population genetic structure of *Bromus tectorum* in the mountains of western North America. *American Journal of Botany*, *104*(6), 879–890.

Assani, A. A., Landry, R., Azouaoui, O., Massicotte, P., & Gratton, D. (2016). Comparison of the characteristics (frequency and timing) of drought and wetness indices of annual mean water levels in the five north american great lakes. *Water Resources Management*, *30*(1), 359–373.

Atyeo, C., & Thackway, R. (2009). Mapping and monitoring revegetation activities in Australia – towards national core attributes. *Australasian Journal of Environmental Management*, *16*(3), 140–148.

Babitha, K. C., Vemanna, R. S., Nataraja, K. N., & Udayakumar, M. (2015). Overexpression of EcbHLH57 transcription factor from *Eleusine coracana L*. in tobacco confers tolerance to salt, oxidative and drought Stress. *PloS One*, *10*(9), e0137098.

Baetscher, D. S., Clemento, A. J., Ng, T. C., Anderson, E. C., & Garza, J. C. (2018). Microhaplotypes provide increased power from short-read DNA sequences for relationship inference. *Molecular Ecology Resources*, *18*(2), 296–305.

Bakker, J. P., & Berendse, F. (1999). Constraints in the restoration of ecological diversity in grassland and heathland communities. *Trends in Ecology & Evolution*. https://www.sciencedirect.com/science/article/pii/S0169534798015444

Bandi, V., & Gutwin, C. (2020). *Interactive Exploration of Genomic Conservation*. https://openreview.net/pdf?id=7-C5VJWbnI

Barluenga, M., Austerlitz, F., Elzinga, J. A., Teixeira, S., Goudet, J., & Bernasconi, G. (2011). Fine-scale spatial genetic structure and gene dispersal in *Silene latifolia*. *Heredity*, *106*(1), 13–24.

Bartoń, K. (2020). *MuMIn: multi-model inference. R package v. 1.43. 6*.

Bastin, J.-F., Finegold, Y., Garcia, C., Mollicone, D., Rezende, M., Routh, D., Zohner, C. M., & Crowther, T. W. (2019). The global tree restoration potential. *Science*, *365*(6448), 76–79.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. In *arXiv [stat.CO]*. arXiv. http://arxiv.org/abs/1406.5823

Baughman, O. W., Agneray, A. C., Forister, M. L., Kilkenny, F. F., Espeland, E. K., Fiegener, R., Horning, M. E., Johnson, R. C., Kaye, T. N., Ott, J., St Clair, J. B., & Leger, E. A. (2019). Strong patterns of intraspecific variation and local adaptation in Great Basin plants revealed through a review of 75 years of experiments. *Ecology and Evolution*, *9*(11), 6259–6275.

Bay, R. A., Harrigan, R. J., Underwood, V. L., Gibbs, H. L., Smith, T. B., & Ruegg, K. (2018). Genomic signals of selection predict climate-driven population declines in a migratory bird. *Science*, *359*(6371), 83–86.

Beierkuhnlein, C., Thiel, D., Jentsch, A., Willner, E., & Kreyling, J. (2011). Ecotypes of European grass species respond differently to warming and extreme drought. *The Journal of Ecology*, *99*(3), 703–713.

Belbin, L., Wallis, E., Hobern, D., & Zerger, A. (2021). The Atlas of Living Australia: History, current state and future directions. *Biodiversity Data Journal*, *9*, e65023.

Bengtsson, J., Bullock, J. M., Egoh, B., Everson, C., Everson, T., O'Connor, T., O'Farrell, P. J., Smith, H. G., & Lindborg, R. (2019). Grasslands-more important for ecosystem services than you might think. In *Ecosphere* (Vol. 10, Issue 2, p. e02582). https://doi.org/10.1002/ecs2.2582

Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., & Huala, E. (2015). The *Arabidopsis* information resource: making and mining the "gold standard" annotated reference plant genome. *Genesis* , *53*(8), 474–485.

Bingham, E. T., Groose, R. W., Woodfield, D. R., & Kidwell, K. K. (1994). Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. *Crop Science*, *34*(4), 823–829.

Birchler, J. A., & Veitia, R. A. (2010). The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution. *The New Phytologist*, *186*(1), 54–62.

Birchler, J. A., & Veitia, R. A. (2012). Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(37), 14746–14753.

Bird, K. A., VanBuren, R., Puzey, J. R., & Edger, P. P. (2018). The causes and consequences of subgenome dominance in hybrids and recent polyploids. *The New Phytologist*, *220*(1), 87–93.

Bittencourt, J. V. M., & Sebbenn, A. M. (2007). Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree *Araucaria angustifolia* in southern Brazil. *Heredity*, *99*(6), 580–591.

Blanc, G., & Wolfe, K. H. (2004). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *The Plant Cell*, *16*(7), 1679–1691.

Blanquart, F., Gandon, S., & Nuismer, S. L. (2012). The effects of migration and drift on local adaptation to a heterogeneous environment. *Journal of Evolutionary Biology*, *25*(7), 1351–1363.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* , *30*(15), 2114–2120.

Booker, T. R., Yeaman, S., & Whitlock, M. C. (2021). The WZA: A window-based method for characterizing genotype-environment association. In *bioRxiv* (p. 2021.06.25.449972). https://doi.org/10.1101/2021.06.25.449972.

Booth, T. H. (2022). Checking bioclimatic variables that combine temperature and precipitation data before their use in species distribution models. *Austral Ecology*, *47*(7), 1506–1514.

Borgström, S., Zachrisson, A., & Eckerberg, K. (2016). Funding ecological restoration policy in practice—patterns of short-termism and regional biases. In *Land Use Policy* (Vol. 52, pp. 439–453). https://doi.org/10.1016/j.landusepol.2016.01.004

Borrell, J. S., Zohren, J., Nichols, R. A., & Buggs, R. J. A. (2020). Genomic assessment of local adaptation in dwarf birch to inform assisted gene flow. *Evolutionary Applications*, *13*(1), 161–175.

Bouckaert, R. R., & Drummond, A. J. (2017). bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evolutionary Biology*, *17*(1), 42.

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., … Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *15*(4), e1006650.

Bradburd, G. S., Ralph, P. L., & Coop, G. M. (2013). Disentangling the effects of geographic and ecological isolation on genetic differentiation. *Evolution; International Journal of Organic Evolution*, *67*(11), 3258–3273.

Brauer, C. J., Hammer, M. P., & Beheregaray, L. B. (2016). Riverscape genomics of a threatened fish across a hydroclimatically heterogeneous river basin. *Molecular Ecology*, *25*(20), 5093–5113.

Breed, M. F., Harrison, P. A., Blyth, C., Byrne, M., Gaget, V., Gellie, N. J. C., Groom, S. V. C., Hodgson, R., Mills, J. G., Prowse, T. A. A., Steane, D. A., & Mohr, J. J. (2019). The potential of genomics for restoring ecosystems and biodiversity. *Nature Reviews. Genetics*, *20*(10), 615–628.

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), 1–10.

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. In *Evolutionary Applications* (Vol. 1, Issue 4, pp. 587–597). https://doi.org/10.1111/j.1752-4571.2008.00045.x

Browning, S. R., & Browning, B. L. (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics*, *81*(5), 1084–1097.

Brown, J. L., Weber, J. J., Alvarado-Serrano, D. F., Hickerson, M. J., Franks, S. J., & Carnaval, A. C. (2016). Predicting the genetic consequences of future climate change: The power of coupling spatial demography, the coalescent, and historical landscape changes. *American Journal of Botany*, *103*(1), 153–163.

Brugnoli, E. A., Urbani, M. H., Quarin, C. L., Zilli, A. L., Martínez, E. J., & Acuña, C. A. (2014). Diversity in apomictic populations of *Paspalum simplex* Morong. *Crop Science*, *54*(4), 1656–1664.

Bucharova, A. (2017). Assisted migration within species range ignores biotic interactions and lacks evidence. *Restoration Ecology*, *25*(1), 14–18.

Buckler, E. S., 4th, Thornsberry, J. M., & Kresovich, S. (2001). Molecular diversity, structure and domestication of grasses. *Genetical Research*, *77*(3), 213–218.

Burnham, K. P., & Anderson, D. R. (2007). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer Science & Business Media.

Bush, A., Mokany, K., Catullo, R., Hoffmann, A., Kellermann, V., Sgrò, C., McEvey, S., & Ferrier, S. (2016). Incorporating evolutionary adaptation in species distribution modelling reduces projected vulnerability to climate change. *Ecology Letters*, *19*(12), 1468–1478.

Bustos-Salazar, A., Smith-Ramírez, C., Zúñiga-Feest, A., Alves, F., & Ivanovich, R. (2017). Which seed origin provides better tolerance to flooding and drought when restoring to face climate change? *Austral Ecology*, *42*(8), 934–946.

Buza, L., Young, A., & Thrall, P. (2000). Genetic erosion, inbreeding and reduced fitness in fragmented populations of the endangered tetraploid pea *Swainsona recta*. *Biological Conservation*, *93*(2), 177–186.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*, 421.

Campbell, M. S., Law, M., Holt, C., Stein, J. C., Moghe, G. D., Hufnagel, D. E., Lei, J., Achawanantakun, R., Jiao, D., Lawrence, C. J., Ware, D., Shiu, S.-H., Childs, K. L., Sun, Y., Jiang, N., & Yandell, M. (2014). MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. *Plant Physiology*, *164*(2), 513–524.

Camus-Kulandaivelu, L., Veyrieras, J.-B., Madur, D., Combes, V., Fourmann, M., Barraud, S., Dubreuil, P., Gouesnard, B., Manicacci, D., & Charcosset, A. (2006). Maize adaptation to temperate climate: relationship between population structure and polymorphism in the Dwarf8 gene. *Genetics*, *172*(4), 2449–2463.

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., Holt, C., Sánchez Alvarado, A., & Yandell, M. (2008). MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, *18*(1), 188–196.

Capblancq, T., Fitzpatrick, M. C., Bay, R. A., Exposito-Alonso, M., & Keller, S. R. (2020). Genomic prediction of (mal)adaptation across current and future climatic landscapes. In *Annual Review of Ecology, Evolution, and Systematics* (Vol. 51, Issue 1, pp. 245–269). https://doi.org/10.1146/annurev-ecolsys-020720-042553

Capblancq, T., Luu, K., Blum, M. G. B., & Bazin, E. (2018). Evaluation of redundancy analysis to identify signatures of local adaptation. *Molecular Ecology Resources*, *18*(6), 1223–1233.

Card DC. Genestats. Retrieved October 2, 2021, from https://github.com/darencard/GenomeAnnotation/blob/97fa52d13eb7a8f6b59a0f6e7261e1e08e542126/genestats

Carretero-Paulet, L., & Van de Peer, Y. (2020). The evolutionary conundrum of whole-genome duplication. *American Journal of Botany*, *107*(8), 1101–1105.

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., Gastauer, M., & Jaffé, R. (2021). Combining genotype, phenotype, and environmental data to delineate site-adjusted provenance strategies for ecological restoration. *Molecular Ecology Resources*, *21*(1), 44–58.

Castellanos-Acuña, D., Vance-Borland, K. W., St. Clair, J. B., Hamann, A., López-Upton, J., Gómez-Pineda, E., Ortega-Rodríguez, J. M., & Sáenz-Romero, C. (2018). Climate-based seed zones for Mexico: guiding reforestation under observed and projected climate change. *New Forests*, *49*(3), 297–309.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140.

Catford, J. A., Bode, M., & Tilman, D. (2018). Introduced species that overcome life history tradeoffs can cause native extinctions. *Nature Communications*, *9*(1), 2131.

Cheng, F., Wu, J., Cai, X., Liang, J., Freeling, M., & Wang, X. (2018). Gene retention, fractionation and subgenome differences in polyploid plants. *Nature Plants*, *4*(5), 258–268.

Cheng, F., Wu, J., Fang, L., Sun, S., Liu, B., Lin, K., Bonnema, G., & Wang, X. (2012). Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*. *PloS One*, *7*(5), e36442.

Chen, Z. J., Sreedasyam, A., Ando, A., Song, Q., De Santiago, L. M., Hulse-Kemp, A. M., Ding, M., Ye, W., Kirkbride, R. C., Jenkins, J., Plott, C., Lovell, J., Lin, Y.-M., Vaughn, R., Liu, B., Simpson, S., Scheffler, B. E., Wen, L., Saski, C. A., … Schmutz, J. (2020). Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nature Genetics*, *52*(5), 525–533.

Cheplick, G. P. (1998). *Population Biology of Grasses*. Cambridge University Press.

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

Clair, J. B. S., St. Clair, J. B., Kilkenny, F. F., Johnson, R. C., Shaw, N. L., & Weaver, G. (2013). Genetic variation in adaptive traits and seed transfer zones for *P seudoroegneria spicata* (bluebunch wheatgrass) in the northwestern U nited S tates. In *Evolutionary Applications* (Vol. 6, Issue 6, pp. 933–948). https://doi.org/10.1111/eva.12077

Clausen, J., Keck, D. D., Hiesey, W. M., & Others. (1940). Experimental studies on the nature of species. Effect of varied environments on western North American plants. *Experimental Studies on the Nature of Species.* https://www.cabdirect.org/cabdirect/abstract/19420702004

Cleverly, J., Eamus, D., Edwards, W., Grant, M., Grundy, M. J., Held, A., Karan, M., Lowe, A. J., Prober, S. M., Sparrow, B., & Others. (2019). TERN, Australia's land observatory: addressing the global challenge of forecasting ecosystem responses to climate variability and change. *Environmental Research Letters: ERL14*(9), 095004.

Comas, L. H., Becker, S. R., Cruz, V. M. V., Byrne, P. F., & Dierig, D. A. (2013). Root traits contributing to plant productivity under drought. *Frontiers in Plant Science*, *4*, 442.

Conant, G. C. (2014). Comparative genomics as a time machine: how relative gene dosage and metabolic requirements shaped the time-dependent resolution of yeast polyploidy. *Molecular Biology and Evolution*, *31*(12), 3184–3193.

Conner, J. K., Hartl, D. L., & Others. (2004). A primer of ecological genetics (Vol. 425). Sinauer Associates Sunderland, MA.

Connor, H. E. (1979). Breeding systems in the grasses: a survey. *New Zealand Journal of Botany*, *17*(4), 547–574.

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, *185*(4), 1411–1423.

Crooks, K. R., & Sanjayan, M. (2006). *Connectivity Conservation*. Cambridge University Press.

Culley, T. M., & Wolfe, A. D. (2001). Population genetic structure of the cleistogamous plant species *Viola pubescens* Aiton (Violaceae), as indicated by allozyme and ISSR molecular markers. *Heredity*, *86*(Pt 5), 545–556.

Cunningham, S. (2008). Rewealth!: stake your claim in the $2 trillion development trend that's renewing the world. McGraw Hill Professional.

Dalgleish, H. J., Koons, D. N., & Adler, P. B. (2010). Can life-history traits predict the response of forb populations to changes in climate variability? *The Journal of Ecology*, *98*(1), 209–217.

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., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics , 27*(15), 2156–2158.

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews. Genetics*, *12*(7), 499–510.

David, P., Pujol, B., Viard, F., Castella, V., & Goudet, J. (2007). Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology*, *16*(12), 2474–2487.

Dawson, T. P., Jackson, S. T., House, J. I., Prentice, I. C., & Mace, G. M. (2011). Beyond predictions: biodiversity conservation in a changing climate. *Science*, *332*(6025), 53–58.

de Faria, A. P., Fernandes, G. W., & França, M. G. C. (2015). Predicting the impact of increasing carbon dioxide concentration and temperature on seed germination and seedling establishment of African grasses in Brazilian Cerrado. *Austral Ecology*, *40*(8), 962–973.

Defoort, J., Van de Peer, Y., & Carretero-Paulet, L. (2019). The evolution of gene duplicates in angiosperms and the impact of protein–protein interactions and the mechanism of duplication. *Genome Biology and Evolution*, *11*(8), 2292–2305.

De Kort, H., Vandepitte, K., & Honnay, O. (2013). A meta-analysis of the effects of plant traits and geographical scale on the magnitude of adaptive differentiation as measured by the difference between QST and FST. *Evolutionary Ecology*, *27*(6), 1081–1097.

Dell'Acqua, M., Fricano, A., Gomarasca, S., Caccianiga, M., Piffanelli, P., Bocchi, S., & Gianfranceschi, L. (2014). Genome scan of Kenyan Themeda triandra populations by AFLP markers reveals a complex genetic structure and hints for ongoing environmental selection. *South African Journal of Botany: Official Journal of the South African Association of Botanists*, *92*, 28–38.

Dengler, J., Janišová, M., Török, P., & Wellstein, C. (2014). Biodiversity of Palaearctic grasslands: a synthesis. *Agriculture, Ecosystems & Environment*, *182*, 1–14.

DeSilva, R., & Dodd, R. S. (2020). Fragmented and isolated: limited gene flow coupled with weak isolation by environment in the paleoendemic giant sequoia (*Sequoiadendron giganteum*). *American Journal of Botany*, *107*(1), 45–55.

De Smet, R., Adams, K. L., Vandepoele, K., Van Montagu, M. C. E., Maere, S., & Van de Peer, Y. (2013). Convergent gene loss following gene and genome duplications creates single-copy families in flowering plants. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(8), 2898–2903.

De Wet, J. M. J., & Harlan, J. R. (1966). Morphology of the compilospecies *Bothriochloa intermedia*. *American Journal of Botany*, *53*(1), 94–98.

De Wet, J. M. J., & Higgins, M. L. (1963). Species relationships within the *Bothriochloa pertusa* complex. *Phyton, Vicente Lopez*. https://www.cabdirect.org/cabdirect/abstract/19651600595

Dfamscan.pl. Retrieved December 10, 2022, from https://dfam.org/releases/current/infrastructure/dfamscan.pl.gz

Dionne, M., Caron, F., Dodson, J. J., & Bernatchez, L. (2008). Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. *Molecular Ecology*, *17*(10), 2382–2396.

Dolezel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, *2*(9), 2233–2244.

Doyle, J. J., Flagel, L. E., Paterson, A. H., Rapp, R. A., Soltis, D. E., Soltis, P. S., & Wendel, J. F. (2008). Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics*, *42*, 443–461.

Dray, S., & Dufour, A.-B. (2007). The ade4 Package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*, 1–20.

Duarte, J. M., Wall, P. K., Edger, P. P., Landherr, L. L., Ma, H., Pires, J. C., Leebens-Mack, J., & dePamphilis, C. W. (2010). Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *BMC Evolutionary Biology*, *10*, 61.

Duminil, J., Hardy, O. J., & Petit, R. J. (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evolutionary Biology*, *9*, 177.

Dunlop, M., Hilbert, D. W., Ferrier, S., House, A., Liedloff, A., Prober, S. M., Smyth, A., Martin, T. G., Harwood, T., Williams, K. J., & Fletcher, C. (2012). The implications of climate change for biodiversity conservation and the National Reserve System: Final synthesis. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.392.7480&rep=rep1&type=pdf

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. *The Journal of Applied Ecology*, *54*(1), 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-Marín, M., & Winn, A. A. (2010). Plant mating systems in a changing world. *Trends in Ecology & Evolution*, *25*(1), 35–43.

Eddy, S. R. (2009). A new generation of homology search tools based on probabilistic inference. *Genome Informatics. International Conference on Genome Informatics*, *23*(1), 205–211.

Edelaar, P., & Björklund, M. (2011). If F(ST) does not measure neutral genetic differentiation, then comparing it with Q(ST) is misleading. Or is it? *Molecular Ecology*, *20*(9), 1805–1812.

Edelaar, P., Burraco, P., & Gomez-Mestre, I. (2011). Comparisons between Q(ST) and F(ST) --how wrong have we been? *Molecular Ecology*, *20*(23), 4830–4839.

Edge, P., Bafna, V., & Bansal, V. (2017). HapCUT2: robust and accurate haplotype assembly for diverse sequencing technologies. *Genome Research*, *27*(5), 801–812.

Edger, P. P., Heidel-Fischer, H. M., Bekaert, M., Rota, J., Glöckner, G., Platts, A. E., Heckel, D. G., Der, J. P., Wafula, E. K., Tang, M., Hofberger, J. A., Smithson, A., Hall, J. C., Blanchette, M., Bureau, T. E., Wright, S. I., dePamphilis, C. W., Eric Schranz, M., Barker, M. S., … Wheat, C. W. (2015). The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(27), 8362–8366.

Edger, P. P., Poorten, T. J., VanBuren, R., Hardigan, M. A., Colle, M., McKain, M. R., Smith, R. D., Teresi, S. J., Nelson, A. D. L., Wai, C. M., Alger, E. I., Bird, K. A., Yocca, A. E., Pumplin, N., Ou, S., Ben-Zvi, G., Brodt, A., Baruch, K., Swale, T., … Knapp, S. J. (2019). Origin and evolution of the octoploid strawberry genome. *Nature Genetics*, *51*(3), 541–547.

Edger, P. P., Smith, R., McKain, M. R., Cooley, A. M., Vallejo-Marin, M., Yuan, Y., Bewick, A. J., Ji, L., Platts, A. E., Bowman, M. J., Childs, K. L., Washburn, J. D., Schmitz, R. J., Smith, G. D., Pires, J. C., & Puzey, J. R. (2017). Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established neo-allopolyploid monkeyflower. *The Plant Cell*, *29*(9), 2150–2167.

Edmands, S. (2007). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology*, *16*(3), 463–475.

Edwards, E. J., Osborne, C. P., Strömberg, C. A. E., Smith, S. A., C4 Grasses Consortium, Bond, W. J., Christin, P.-A., Cousins, A. B., Duvall, M. R., Fox, D. L., Freckleton, R. P., Ghannoum, O., Hartwell, J., Huang, Y., Janis, C. M., Keeley, J. E., Kellogg, E. A., Knapp, A. K., Leakey, A. D. B., … Tipple, B. (2010). The origins of C4 grasslands: integrating evolutionary and ecosystem science. *Science*, *328*(5978), 587–591.

Ehrendorfer, F. (1979). Polyploidy and distribution. *Basic Life Sciences*, *13*, 45–60.

Ellinghaus, D., Kurtz, S., & Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics*, *9*, 18.

Ellis, N. R., & Albrecht, G. A. (2017). Climate change threats to family farmers' sense of place and mental wellbeing: A case study from the Western Australian Wheatbelt. *Social Science & Medicine*, *175*, 161–168.

Ellis, N., Smith, S. J., & Pitcher, C. R. (2012). Gradient forests: calculating importance gradients on physical predictors. *Ecology*, *93*(1), 156–168.

Ellstrand, N. C., & Elam, D. R. (1993). Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics*, *24*, 217–242.

Emery, M., Willis, M. M. S., Hao, Y., Barry, K., Oakgrove, K., Peng, Y., Schmutz, J., Lyons, E., Pires, J. C., Edger, P. P., & Conant, G. C. (2018). Preferential retention of genes from one parental genome after polyploidy illustrates the nature and scope of the genomic conflicts induced by hybridization. *PLoS Genetics*, *14*(3), e1007267.

Emms, D. M., & Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1), 238.

Engelbrecht, B. M. J., Dalling, J. W., Pearson, T. R. H., Wolf, R. L., Gálvez, D. A., Koehler, T., Tyree, M. T., & Kursar, T. A. (2006). Short dry spells in the wet season increase mortality of tropical pioneer seedlings. *Oecologia*, *148*(2), 258–269.

Environment protection and biodiversity conservation act 1999.

Estep, M. C., McKain, M. R., Vela Diaz, D., Zhong, J., Hodge, J. G., Hodkinson, T. R., Layton, D. J., Malcomber, S. T., Pasquet, R., & Kellogg, E. A. (2014). Allopolyploidy, diversification, and the Miocene grassland expansion. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(42), 15149–15154.

Etterson, J. R. (2004). Evolutionary potential of *Chamaecrista fasciculata* in relation to climate change. I. Clinal patterns of selection along an environmental gradient in the great plains. *Evolution; International Journal of Organic Evolution*, *58*(7), 1446–1458.

Etterson, J. R., Keller, S. R., & Galloway, L. F. (2007). Epistatic and cytonuclear interactions govern outbreeding depression in the autotetraploid *Campanulastrum americanum*. *Evolution; International Journal of Organic Evolution*, *61*(11), 2671–2683.

Exposito-Alonso, M., Vasseur, F., Ding, W., Wang, G., Burbano, H. A., & Weigel, D. (2018). Genomic basis and evolutionary potential for extreme drought adaptation in *Arabidopsis thaliana*. *Nature Ecology & Evolution*, *2*(2), 352–358.

Falconer, & MacKay. (1996). Introduction to quantitative genetics. Burnt Mill. *Harlow, UK: Longman Scientific & Technical*.

Faske, T. M., Agneray, A. C., Jahner, J. P., Sheta, L. M., Leger, E. A., & Parchman, T. L. (2021). Genomic and common garden approaches yield complementary results for quantifying

environmental drivers of local adaptation in rubber rabbitbrush, a foundational Great Basin shrub. *Evolutionary Applications*, *14*(12), 2881–2900.

Fay, P. A., & Schultz, M. J. (2009). Germination, survival, and growth of grass and forb seedlings: Effects of soil moisture variability. *Acta Oecologica*, *35*(5), 679–684.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315.

Fiedler, S., Perring, M. P., & Tietjen, B. (2018). Integrating trait-based empirical and modeling research to improve ecological restoration. *Ecology and Evolution*, *8*(12), 6369–6380.

Fisher, R. A. (1919). XV.—The correlation between relatives on the supposition of mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, *52*(2), 399–433.

Fitter, A. H., & Silvertown, J. W. (1983). Introduction to plant population ecology. In *The Journal of Applied Ecology* (Vol. 20, Issue 2, p. 693). https://doi.org/10.2307/2403548

Fitzpatrick, M. C., Chhatre, V. E., Soolanayakanahally, R. Y., & Keller, S. R. (2021). Experimental support for genomic prediction of climate maladaptation using the machine learning approach Gradient Forests. *Molecular Ecology Resources*, *21*(8), 2749–2765.

Fitzpatrick, M. C., & Keller, S. R. (2015). Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation. *Ecology Letters*, *18*(1), 1–16.

Flagel, L. E., & Wendel, J. F. (2009). Gene duplication and evolutionary novelty in plants. *The New Phytologist*, *183*(3), 557–564.

Flagel, L. E., & Wendel, J. F. (2010). Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. *The New Phytologist*, *186*(1), 184–193.

Folk, R. A., Siniscalchi, C. M., & Soltis, D. E. (2020). Angiosperms at the edge: Extremity, diversity, and phylogeny. *Plant, Cell & Environment*, *43*(12), 2871–2893.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, *151*(4), 1531–1545.

Forester, B. R., Jones, M. R., Joost, S., Landguth, E. L., & Lasky, J. R. (2016). Detecting spatial genetic signatures of local adaptation in heterogeneous landscapes. *Molecular Ecology*, *25*(1), 104–120.

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype-environment associations. *Molecular Ecology*, *27*(9), 2215–2233.

Fowler, N. L., & Levin, D. A. (1984). Ecological constraints on the establishment of a novel polyploid in competition with its diploid progenitor. *The American Naturalist*, *124*(5), 703–711.

Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, *17*(1), 27–32.

Frankham, R., Ballou, J. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology: The Journal of the Society for Conservation Biology*, *25*(3), 465–475.

Freeling, M. (2009). Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition. *Annual Review of Plant Biology*, *60*, 433–453.

Freeling, M., Woodhouse, M. R., Subramaniam, S., Turco, G., Lisch, D., & Schnable, J. C. (2012). Fractionation mutagenesis and similar consequences of mechanisms removing dispensable or less-expressed DNA in plants. *Current Opinion in Plant Biology*, *15*(2), 131–139.

Friedman, J., & Barrett, S. C. H. (2009). Wind of change: new insights on the ecology and evolution of pollination and mating in wind-pollinated plants. *Annals of Botany*, *103*(9), 1515–1527.

Fuentes-Pardo, A. P., & Ruzzante, D. E. (2017). Whole-genome sequencing approaches for conservation biology: Advantages, limitations and practical recommendations. *Molecular Ecology*, *26*(20), 5369–5406.

Futuyma, D. J. (2013). The evolution of evolutionary ecology. *Israel Journal of Ecology & Evolution*, *59*(4), 172–180.

Galbraith, D. W., & Lambert, G. M. (2012). High-throughput monitoring of plant nuclear DNA contents via flow cytometry. *Methods in Molecular Biology* , *918*, 311–325.

Galliart, M., Bello, N., Knapp, M., Poland, J., St Amand, P., Baer, S., Maricle, B., Smith, A. B., & Johnson, L. (2019). Local adaptation, genetic divergence, and experimental selection in a foundation grass across the US Great Plains' climate gradient. In *Global Change Biology* (Vol. 25, Issue 3, pp. 850–868). https://doi.org/10.1111/gcb.14534

Galloway, L. F., & Fenster, C. B. (2000). Population differentiation in an annual legume: local adaptation. *Evolution; International Journal of Organic Evolution*, *54*(4), 1173–1181.

García-Dorado, A. (2012). Understanding and predicting the fitness decline of shrunk populations: inbreeding, purging, mutation, and standard selection. *Genetics*, *190*(4), 1461–1476.

Garnier, S., Ross, N., Rudis, R., Camargo, P. A., Sciaini, M., & Scherer, C. (2021). Viridis—Colorblind-Friendly Color Maps for R. *R Package Version 0. 6*, *2*.

Garsmeur, O., Schnable, J. C., Almeida, A., Jourda, C., D'Hont, A., & Freeling, M. (2014). Two evolutionarily distinct classes of paleopolyploidy. *Molecular Biology and Evolution*, *31*(2), 448–454.

Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*, *201*(4), 1555–1579.

Gautier, M., Yamaguchi, J., Foucaud, J., Loiseau, A., Ausset, A., Facon, B., Gschloessl, B., Lagnel, J., Loire, E., Parrinello, H., Severac, D., Lopez-Roques, C., Donnadieu, C., Manno, M., Berges, H., Gharbi, K., Lawson-Handley, L., Zang, L.-S., Vogel, H., … Prud'homme, B. (2018). The genomic basis of color pattern polymorphism in the harlequin ladybird. *Current Biology: CB*, *28*(20), 3296–3302.e7.

Gellesch, E., Arfin Khan, M. A. S., Jentsch, A., & Beierkuhnlein, C. (2017). Grassland experiments under climatic extremes: Reproductive fitness versus biomass. *Environmental and Experimental Botany*, *144*, 68–75.

Gergis, J., & Ashcroft, L. (2013). Rainfall variations in south-eastern Australia part 2: a comparison of documentary, early instrumental and palaeoclimate records, 1788-2008. *International Journal of Climatology*, *33*(14), 2973–2987.

Ghazoul, J. (2005). Pollen and seed dispersal among dispersed plants. *Biological Reviews of the Cambridge Philosophical Society*, *80*(3), 413–443.

Ghildiyal, S. K. (2009). Environmental variation in seed and seedling characteristics of *pinus roxburghii* sarg. from uttarakhand, india. In *Applied Ecology and Environmental Research* (Vol. 7, Issue 2, pp. 121–129). https://doi.org/10.15666/aeer/0702_121129

Gibbons, P., & Boak, M. (2002). The value of paddock trees for regional conservation in an agricultural landscape. *Ecological Management & Restoration*, *3*(3), 205–210.

Gibbons, P., Briggs, S. V., Murphy, D. Y., Lindenmayer, D. B., McElhinny, C., & Brookhouse, M. (2010). Benchmark stem densities for forests and woodlands in south-eastern Australia under conditions of relatively little modification by humans since European settlement. *Forest Ecology and Management*, *260*(12), 2125–2133.

Gilbert, G. S., Harms, K. E., Hamill, D. N., & Hubbell, S. P. (2001). Effects of seedling size, El Niño drought, seedling density, and distance to nearest conspecific adult on 6-year survival of *Ocotea whitei* seedlings in Panamá. *Oecologia*, *127*(4), 509–516.

Gilbert, K. J., & Whitlock, M. C. (2015). QST-FST comparisons with unbalanced half-sib designs. *Molecular Ecology Resources*, *15*(2), 262–267.

Gleaves, J. T. (1973). Gene flow mediated by wind-borne pollen. *Heredity*, *31*(3), 355–366.

Godfree, R. C., Marshall, D. J., Young, A. G., Miller, C. H., & Mathews, S. (2017). Empirical evidence of fixed and homeostatic patterns of polyploid advantage in a keystone grass exposed to drought and heat stress. *Royal Society Open Science*, *4*(11), 170934.

Godt, M. J. W., & Hamrick, J. L. (1998). Allozyme diversity in the grasses. In *Population Biology of Grasses* (pp. 11–29). https://doi.org/10.1017/cbo9780511525445.003

Gomulkiewicz, R., Holt, R. D., & Barfield, M. (1999). The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. *Theoretical Population Biology*, *55*(3), 283–296.

González Barrios, P., Speranza, P., Glison, N., Piccardi, M., Balzarini, M., & Gutiérrez, L. (2016). Analysis of flowering dynamics heritability in the perennial warm-season grass Paspalum dilatatum. *Grass and Forage Science*, *71*(1), 123–131.

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research*, *40*(Database issue), D1178–D1186.

Goodwillie, C., Kalisz, S., & Eckert, C. G. (2005). The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics*, *36*(1), 47–79.

Goslee, S. C., & Urban, D. L. (2007). The ecodist Package for Dissimilarity-based Analysis of Ecological Data. *Journal of Statistical Software*, *22*, 1–19.

Gould, B. A., Palacio-Mejia, J. D., Jenkins, J., Mamidi, S., Barry, K., Schmutz, J., Juenger, T. E., & Lowry, D. B. (2018). Population genomics and climate adaptation of a C4 perennial grass, *Panicum hallii* (Poaceae). *BMC Genomics*, *19*(1), 792.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652.

Grainger, A. (1994). Book reviews : United Nations Environment Programme. 1992: World atlas of desertification. Sevenoaks: Edward Arnold. x 70 pp. £85.00 cloth. ISBN 0 340 555122. In *Progress in Physical Geography: Earth and Environment* (Vol. 18, Issue 4, pp. 621–623). https://doi.org/10.1177/030913339401800420

Grant, V. (1981). Plant Speciation. https://doi.org/10.7312/gran92318

Gray, M. M., St Amand, P., Bello, N. M., Galliart, M. B., Knapp, M., Garrett, K. A., Morgan, T. J., Baer, S. G., Maricle, B. R., Akhunov, E. D., & Johnson, L. C. (2014). Ecotypes of an ecologically dominant prairie grass (*Andropogon gerardii*) exhibit genetic divergence across the U.S. Midwest grasslands' environmental gradient. *Molecular Ecology*, *23*(24), 6011–6028.

Greenville, A. C., Dickman, C. R., & Wardle, G. M. (2017). 75 years of dryland science: Trends and gaps in arid ecology literature. *PloS One*, *12*(4), e0175014.

Griffiths, A. G., Moraga, R., Tausen, M., Gupta, V., Bilton, T. P., Campbell, M. A., Ashby, R., Nagy, I., Khan, A., Larking, A., Anderson, C., Franzmayr, B., Hancock, K., Scott, A., Ellison, N. W., Cox, M.

P., Asp, T., Mailund, T., Schierup, M. H., & Andersen, S. U. (2019). Breaking free: the genomics of allopolyploidy-facilitated niche expansion in white clover. *The Plant Cell*, *31*(7), 1466–1487.

Grubb, P. J. (1977). The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological Reviews of the Cambridge Philosophical Society*, *52*(1), 107–145.

Gruber, B., Unmack, P. J., Berry, O. F., & Georges, A. (2018). dartr : An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. In *Molecular Ecology Resources* (Vol. 18, Issue 3, pp. 691–699). https://doi.org/10.1111/1755-0998.12745

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, *52*(5), 696–704.

Guk, J.-Y., Jang, M.-J., Choi, J.-W., Lee, Y. M., & Kim, S. (2022). De novo phasing resolves haplotype sequences in complex plant genomes. *Plant Biotechnology Journal*, *20*(6), 1031–1041.

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics*, *195*(1), 205–220.

Guo, J., Brown, P. J., Rayburn, A. L., Butts-Wilmsmeyer, C. J., Boe, A., & Lee, D. (2021). Genomic variation shaped by environmental and geographical factors in prairie cordgrass natural populations collected across its native range in the USA. *Genes*, *12*(8). https://doi.org/10.3390/genes12081240

Gurgul, A., Miksza-Cybulska, A., Szmatoła, T., Jasielczuk, I., Piestrzyńska-Kajtoch, A., Fornal, A., Semik-Gurgul, E., & Bugno-Poniewierska, M. (2019). Genotyping-by-sequencing performance in selected livestock species. *Genomics*, *111*(2), 186–195.

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* , *32*(18), 2847–2849.

Hagon, M. W. (1977). Effects of competition, herbicides and activated carbon on establishment of Australian grasses. *Weed Research*, *17*(5), 297–301.

Haig, D. (2020). From Darwin to Derrida: selfish genes, social selves, and the meanings of life. MIT Press.

Haig, I. T., Weidman, R. H., & Davis, K. P. (1941). Natural Regeneration in the Western White Pine Type. The Department.

Hall, D., Luquez, V., Garcia, V. M., St Onge, K. R., Jansson, S., & Ingvarsson, P. K. (2007). Adaptive population differentiation in phenology across a latitudinal gradient in European aspen (*Populus tremula*, L.): a comparison of neutral markers, candidate genes and phenotypic traits. *Evolution; International Journal of Organic Evolution*, *61*(12), 2849–2860.

Hamann, A., Gylander, T., & Chen, P.-Y. (2011). Developing seed zones and transfer guidelines with multivariate regression trees. *Tree Genetics & Genomes*, *7*(2), 399–408.

Hamann, E., Kesselring, H., Armbruster, G. F. J., Scheepens, J. F., & Stöcklin, J. (2016). Evidence of local adaptation to fine- and coarse-grained environmental variability in *Poa alpina* in the Swiss Alps. *The Journal of Ecology*, *104*(6), 1627–1637.

Handel, S. N. (1983). Contrasting gene flow patterns and genetic subdivision in adjacent populations of *cucumis sativus* (cucurbitaceae). *Evolution; International Journal of Organic Evolution*, *37*(4), 760–771.

Han, Y., & Wessler, S. R. (2010). MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Research*, *38*(22), e199.

Harlan, J. R., & de Wet, J. M. J. (1963). The Compilospecies Concept. *Evolution; International Journal of Organic Evolution*, *17*(4), 497–501.

Harper, J. L. (1977). Population biology of plants. *Population Biology of Plants.* https://www.cabdirect.org/cabdirect/abstract/19782321379

Harris, J. A., Hobbs, R. J., Higgs, E., & Aronson, J. (2006). Ecological Restoration and Global Climate Change. *Restoration Ecology*, *14*(2), 170–176.

Harrison, S. P., Gornish, E. S., & Copeland, S. (2015). Climate-driven diversity loss in a grassland community. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(28), 8672–8677.

Heller, N. E., & Zavaleta, E. S. (2009). Biodiversity management in the face of climate change: A review of 22 years of recommendations. *Biological Conservation*, *142*(1), 14–32.

Hendry, A. P. (2013). Key questions in the genetics and genomics of eco-evolutionary dynamics. *Heredity*, *111*(6), 456–466.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, *173*(5), 579–588.

Hewitt, N., Klenk, N., Smith, A. L., Bazely, D. R., Yan, N., Wood, S., MacLellan, J. I., Lipsig-Mumme, C., & Henriques, I. (2011). Taking stock of the assisted migration debate. *Biological Conservation*, *144*(11), 2560–2572.

He, X., & Zhang, J. (2005). Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics*, *169*(2), 1157–1164.

Hijmans, R. J., & van Etten, J. (2012). *raster: Geographic analysis and modeling with raster data. R package version 2.0-12*.

Hijmans, R. J., Williams, E., Vennes, C., & Hijmans, M. R. J. (2017). Package "geosphere." *Spherical Trigonometry*, *1*(7). ftp://sunsite2.icm.edu.pl/site/cran/web/packages/geosphere/geosphere.pdf

Himmelmann, M. L. (2022). *Package "HMM."* https://cran.opencpu.org/web/packages/HMM/HMM.pdf

Hirsch, C. D., & Springer, N. M. (2017). Transposable element influences on gene expression in plants. *Biochimica et Biophysica Acta, Gene Regulatory Mechanisms*, *1860*(1), 157–165.

Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M. L., Reed, L. K., Storfer, A., & Whitlock, M. C. (2016). Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *The American Naturalist*, *188*(4), 379–397.

Hobbs, R. J., & Yates, C. J. (2000). Temperate Eucalypt Woodlands in Australia*: Biology, Conservation, Management and Restoration*. Surrey Beatty & Sons.

Hodgins, K. A., & Moore, J. L. (2016). Adapting to a warming world: Ecological restoration, climate change, and genomics. *American Journal of Botany*, *103*(4), 590–592.

Hoegh-Guldberg, O., Hughes, L., McIntyre, S., Lindenmayer, D. B., Parmesan, C., Possingham, H. P., & Thomas, C. D. (2008). Ecology. Assisted colonization and rapid climate change. *Science*, *321*(5887), 345–346.

Hoffmann, A. A., Rymer, P. D., Byrne, M., Ruthrof, K. X., Whinam, J., McGeoch, M., Bergstrom, D. M., Guerin, G. R., Sparrow, B., Joseph, L., Hill, S. J., Andrew, N. R., Camac, J., Bell, N., Riegler, M., Gardner, J. L., & Williams, S. E. (2019). Impacts of recent climate change on terrestrial flora and fauna: Some emerging Australian examples. *Austral Ecology*, *44*(1), 3–27.

Hoffmann, A. A., & Sgrò, C. M. (2011). Climate change and evolutionary adaptation. *Nature*, *470*(7335), 479–485.

Hoffman, J. I., Simpson, F., David, P., Rijks, J. M., Kuiken, T., Thorne, M. A. S., Lacy, R. C., & Dasmahapatra, K. K. (2014). High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(10), 3775–3780.

Hoffmann, A. A., Weeks, A. R., & Sgrò, C. M. (2021). Opportunities and challenges in assessing climate change vulnerability through genomics. *Cell*, *184*(6), 1420–1425.

Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, *30*(1), 62–82.

Holl, K. D., & Howarth, R. B. (2000). Paying for restoration. *Restoration Ecology*, *8*(3), 260–267.

Honnay, O., & Jacquemyn, H. (2007). Susceptibility of common and rare plant species to the genetic consequences of habitat fragmentation. *Conservation Biology: The Journal of the Society for Conservation Biology*, *21*(3), 823–831.

Huang, C.-L., Chen, J.-H., Chang, C.-T., Chung, J.-D., Liao, P.-C., Wang, J.-C., & Hwang, S.-Y. (2016). Disentangling the effects of isolation-by-distance and isolation-by-environment on genetic differentiation among *Rhododendron* lineages in the subgenus *Tsutsusi*. *Tree Genetics & Genomes*, *12*(3). https://doi.org/10.1007/s11295-016-1010-2

Hu, Y., Burucs, Z., & Schmidhalter, U. (2006). Short-Term Effect of Drought and Salinity on Growth and Mineral Elements in Wheat Seedlings. *Journal of Plant Nutrition*, *29*(12), 2227–2243.

Hudson, A. R., Ayre, D. J., & Ooi, M. K. J. (2015). Physical dormancy in a changing climate. *Seed Science Research*, *25*(2), 66–81.

Hufford, K. M., & Mazer, S. J. (2003). Plant ecotypes: genetic differentiation in the age of ecological restoration. *Trends in Ecology & Evolution*, *18*(3), 147–155.

Hufford, K. M., & Mazer, S. J. (2012). Local adaptation and the effects of grazing on the performance of *Nassella pulchra*: Implications for seed sourcing in restoration. *Restoration Ecology*, *20*(6), 688–695.

Ikeda, D. H., Max, T. L., Allan, G. J., Lau, M. K., Shuster, S. M., & Whitham, T. G. (2017). Genetically informed ecological niche models improve climate change predictions. *Global Change Biology*, *23*(1), 164–176.

Ingvarsson, P. K., & Bernhardsson, C. (2020). Genome-wide signatures of environmental adaptation in European aspen (*Populus tremula*) under current and future climate conditions. *Evolutionary Applications*, *13*(1), 132–142.

Jackson, S. D. (2009). Plant responses to photoperiod. *The New Phytologist*, *181*(3), 517–531.

Janakiram, T., Namita, N., & Others. (2014). Genetic divergence analysis in turf grasses based on morphological traits. *Indian Journal of Agricultural Sciences*, *84*(9), 1035–1039.

Jia, K.-H., Zhao, W., Maier, P. A., Hu, X.-G., Jin, Y., Zhou, S.-S., Jiao, S.-Q., El-Kassaby, Y. A., Wang, T., Wang, X.-R., & Mao, J.-F. (2020). Landscape genomics predicts climate change-related genetic offset for the widespread *Platycladus orientalis* (Cupressaceae). *Evolutionary Applications*, *13*(4), 665–676.

Jiang, N., Bao, Z., Zhang, X., Hirochika, H., Eddy, S. R., McCouch, S. R., & Wessler, S. R. (2003). An active DNA transposon family in rice. *Nature*, *421*(6919), 163–167.

Jiao, H., Wang, Y., Zhang, L., Jiang, P., & Zhao, H. (2018). Lineage-specific duplication and adaptive evolution of bitter taste receptor genes in bats. *Molecular Ecology*, *27*(22), 4475–4488.

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., Tomsho, L. P., Hu, Y., Liang, H., Soltis, P. S., Soltis, D. E., Clifton, S. W., Schlarbaum, S. E., Schuster, S. C., Ma, H., Leebens-Mack, J., & dePamphilis, C. W. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature*, *473*(7345), 97–100.

Johnson, R. N., O'Meally, D., Chen, Z., Etherington, G. J., Ho, S. Y. W., Nash, W. J., Grueber, C. E., Cheng, Y., Whittington, C. M., Dennison, S., Peel, E., Haerty, W., O'Neill, R. J., Colgan, D., Russell, T. L., Alquezar-Planas, D. E., Attenbrow, V., Bragg, J. G., Brandies, P. A., … Belov, K.

(2018). Adaptation and conservation insights from the koala genome. *Nature Genetics*, *50*(8), 1102–1111.

Johnson, W. E., & Koepfli, K. (2014). The role of genomics in conservation and reproductive sciences. *Advances in Experimental Medicine and Biology*, *753*, 71–96.

Jones, A. T., Hayes, M. J., & Sackville Hamilton, N. R. (2001). The effect of provenance on the performance of *Crataegus monogyna* in hedges. *The Journal of Applied Ecology*, *38*(5), 952–962.

Jones, N. T., Husband, B. C., & MacDougall, A. S. (2013). Reproductive system of a mixed-mating plant responds to climate perturbation by increased selfing. *Proceedings. Biological Sciences / The Royal Society*, *280*(1766), 20131336.

Jordan, R., Hoffmann, A. A., Dillon, S. K., & Prober, S. M. (2017). Evidence of genomic adaptation to climate in *Eucalyptus microcarpa*: Implications for adaptive potential to projected climate change. *Molecular Ecology*, *26*(21), 6002–6020.

Jordan, R., Prober, S. M., Hoffmann, A. A., & Dillon, S. K. (2020). Combined analyses of phenotype, genotype and climate implicate local adaptation as a driver of diversity in *Eucalyptus microcarpa* (Grey box). *Forests, Trees and Livelihoods*, *11*(5), 495.

Joshi, J., Schmid, B., Caldeira, M. C., Dimitrakopoulos, P. G., Good, J., Harris, R., Hector, A., Huss-Danell, K., Jumpponen, A., Minns, A., Mulder, C. P. H., Pereira, J. S., Prinz, A., Scherer-Lorenzen, M., Siamantziouras, A.-S. D., Terry, A. C., Troumbis, A. Y., & Lawton, J. H. (2001). Local adaptation enhances performance of common plant species. *Ecology Letters*, *4*(6), 536–544.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ*, *2*, e281.

Kanehisa, M. (2019). Toward understanding the origin and evolution of cellular organisms. *Protein Science: A Publication of the Protein Society*, *28*(11), 1947–1951.

Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M., & Tanabe, M. (2021). KEGG: integrating viruses and cellular organisms. *Nucleic Acids Research*, *49*(D1), D545–D551.

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, *28*(1), 27–30.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, *7*(12), 1225–1241.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An Rpackage for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution / British Ecological Society*, *4*(8), 782–788.

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology & Evolution*, *17*(5), 230–241.

Kennedy, R. C., Unger, M. F., Christley, S., Collins, F. H., & Madey, G. R. (2011). An automated homology-based approach for identifying transposable elements. *BMC Bioinformatics*, *12*, 130.

Khan, S., Nabi, G., Ullah, M. W., Yousaf, M., Manan, S., Siddique, R., & Hou, H. (2016). Overview on the Role of Advance Genomics in Conservation Biology of Endangered Species. *International Journal of Genomics and Proteomics*, *2016*, 3460416.

Kiem, A. S., Johnson, F., Westra, S., van Dijk, A., Evans, J. P., O'Donnell, A., Rouillard, A., Barr, C., Tyler, J., Thyer, M., Jakob, D., Woldemeskel, F., Sivakumar, B., & Mehrotra, R. (2016). Natural hazards in Australia: droughts. *Climatic Change*, *139*(1), 37–54.

King, A. D., Pitman, A. J., Henley, B. J., Ukkola, A. M., & Brown, J. R. (2020). The role of climate variability in Australian drought. *Nature Climate Change*, *10*(3), 177–179.

Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hill, C. E., Hoang, A., Gibert, P., & Beerli, P. (2001). The strength of phenotypic selection in natural populations. *The American Naturalist*, *157*(3), 245–261.

Kirono, D. G. C., Hennessy, K. J., & Grose, M. R. (2017). Increasing risk of months with low rainfall and high temperature in southeast Australia for the past 150 years. *Climate Risk Management*, *16*, 10–21.

Kleine, T., Voigt, C., & Leister, D. (2009). Plastid signalling to the nucleus: messengers still lost in the mists? *Trends in Genetics: TIG*, *25*(4), 185–192.

Knapp, E. E., & Rice, K. J. (1994). Starting from seed genetic issues in using native grasses for restoration. *Ecological Restoration* , *12*(1), 40–45.

Koch, J. M., & Samsa, G. P. (2007). Restoring Jarrah forest trees after bauxite mining in Western Australia. *Restoration Ecology*. https://doi.org/10.1111/j.1526-100X.2007.00289.x

Koelling, V. A., Hamrick, J. L., & Mauricio, R. (2011). Genetic diversity and structure in two species of *Leavenworthia* with self-incompatible and self-compatible populations. *Heredity*, *106*(2), 310–318.

Köhler, C., Mittelsten Scheid, O., & Erilova, A. (2010). The impact of the triploid block on the origin and evolution of polyploid plants. *Trends in Genetics: TIG*, *26*(3), 142–148.

Kolb, T. E., Grady, K. C., McEttrick, M. P., & Herrero, A. (2016). Local-scale drought adaptation of *ponderosa pine* seedlings at habitat ecotones. *Forest Science*, *62*(6), 641–651.

Koren, S., Rhie, A., Walenz, B. P., Dilthey, A. T., Bickhart, D. M., Kingan, S. B., Hiendleder, S., Williams, J. L., Smith, T. P. L., & Phillippy, A. M. (2018). De novo assembly of haplotype-resolved genomes with trio binning. *Nature Biotechnology*. https://doi.org/10.1038/nbt.4277

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics*, *5*, 59.

Kramer, D. L., Maricle, K. L., Hilt, C. J., Martin, N. M., Urban, A. D., Smart, C. M., Baer, S. G., Johnson, L. C., & Maricle, B. R. (2018). Drought tolerance in ecotypes of big bluestem (*Andropogon gerardii*) relates to above-ground surface area: Results from a common garden experiment. *Flora* , *246-247*, 52–60.

Kranabetter, J. M., Stoehr, M. U., & O'Neill, G. A. (2012). Divergence in ectomycorrhizal communities with foreign Douglas-fir populations and implications for assisted migration. *Ecological Applications: A Publication of the Ecological Society of America*, *22*(2), 550–560.

Kreyling, J., Bittner, T., Jaeschke, A., Jentsch, A., Steinbauer, M. J., Thiel, D., & Beierkuhnlein, C. (2011). Assisted colonization: a question of focal units and recipient localities. In *Restoration Ecology* (Vol. 19, Issue 4, pp. 433–440). https://doi.org/10.1111/j.1526-100x.2011.00777.x

Kronenberg, Z. N., Rhie, A., Koren, S., Concepcion, G. T., Peluso, P., Munson, K. M., Porubsky, D., Kuhn, K., Mueller, K. A., Low, W. Y., Hiendleder, S., Fedrigo, O., Liachko, I., Hall, R. J., Phillippy, A. M., Eichler, E. E., Williams, J. L., Smith, T. P. L., Jarvis, E. D., … Kingan, S. B. (2021). Extended haplotype-phasing of long-read de novo genome assemblies using Hi-C. *Nature Communications*, *12*(1), 1935.

Kruuk, L. E. B., Slate, J., & Wilson, A. J. (2008). New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annual Review of Ecology, Evolution, and Systematics*, *39*, 525–548.

Kumar, A., Yau, Y.-Y., Ogita, S., & Scheibe, R. (2020). Climate change, photosynthesis and advanced biofuels: the role of biotechnology in the production of value-added plant bio-products. Springer Nature.

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, *34*(7), 1812–1819.

Kurze, S., Bareither, N., & Metz, J. (2017). Phenology, roots and reproductive allocation, but not the LHS scheme, shape ecotypes along an aridity gradient. *Perspectives in Plant Ecology, Evolution and Systematics*, *29*, 20–29.

Lacey, E. P. (1998). What is an adaptive environmentally induced parental effect. *Maternal Effects as Adaptations*, 54–66.

Lamy, J.-B., Bouffier, L., Burlett, R., Plomion, C., Cochard, H., & Delzon, S. (2011). Uniform selection as a primary force reducing population genetic differentiation of cavitation resistance across a species range. *PloS One*, *6*(8), e23476.

Langlet, O. (1971). Two hundred years genecology. *Taxon*, *20*(5/6), 653–721.

Láruson, Á. J., Fitzpatrick, M. C., Keller, S. R., Haller, B. C., & Lotterhos, K. E. (2022). Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest. *Evolutionary Applications*, *15*(3), 403–416.

Leberg, P. L., & Firmin, B. D. (2008). Role of inbreeding depression and purging in captive breeding and restoration programmes. *Molecular Ecology*, *17*(1), 334–343.

Leck, M. A., Thomas Parker, V., Simpson, R. L., & Simpson, R. S. (2008). Seedling Ecology and Evolution. Cambridge University Press.

Le Corre, V., & Kremer, A. (2012). The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, *21*(7), 1548–1566.

Ledyard Stebbins, G. (1950). *Variation and Evolution in Plants*. Columbia University Press.

Leimu, R., & Fischer, M. (2008). A meta-analysis of local adaptation in plants. *PloS One*, *3*(12), e4010.

Leimu, R., Mutikainen, P., Koricheva, J., & Fischer, M. (2006). How general are positive relationships between plant population size, fitness and genetic variation? *The Journal of Ecology*, *94*(5), 942–952.

Leinonen, T., O'Hara, R. B., Cano, J. M., & Merilä, J. (2008). Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, *21*(1), 1–17.

Leinonen, T., Scott McCairns, R. J., O'Hara, R. B., & Merilä, J. (2013). QST–FST comparisons: evolutionary and ecological insights from genomic heterogeneity. In *Nature Reviews Genetics* (Vol. 14, Issue 3, pp. 179–190). https://doi.org/10.1038/nrg3395

Leroy, T., Louvet, J.-M., Lalanne, C., Le Provost, G., Labadie, K., Aury, J.-M., Delzon, S., Plomion, C., & Kremer, A. (2020). Adaptive introgression as a driver of local adaptation to climate in European white oaks. *The New Phytologist*, *226*(4), 1171–1182.

Levin, D. A. (1975). Minority cytotype exclusion in local plant populations. *Taxon*, *24*(1), 35–43.

Levin, D. A. (1983). Polyploidy and novelty in flowering plants. *The American Naturalist*, *122*(1), 1–25.

Li, C., Yan, C., Sun, Q., Wang, J., Yuan, C., Mou, Y., Shan, S., & Zhao, X. (2021). The bHLH transcription factor AhbHLH112 improves the drought tolerance of peanut. *BMC Plant Biology*, *21*(1), 540.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, *326*(5950), 289–293.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* , *34*(18), 3094–3100.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* , *25*(14), 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* , *25*(16), 2078–2079.

Lim, K. Y., Soltis, D. E., Soltis, P. S., Tate, J., Matyasek, R., Srubarova, H., Kovarik, A., Pires, J. C., Xiong, Z., & Leitch, A. R. (2008). Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae). *PloS One*, *3*(10), e3353.

Linder, H. P., & Barker, N. P. (2014). Does polyploidy facilitate long-distance dispersal? *Annals of Botany*, *113*(7), 1175–1183.

Linder, H. P., Lehmann, C. E. R., Archibald, S., Osborne, C. P., & Richardson, D. M. (2018). Global grass (Poaceae) success underpinned by traits facilitating colonization, persistence and habitat transformation. In *Biological Reviews* (Vol. 93, Issue 2, pp. 1125–1144). https://doi.org/10.1111/brv.12388

Linhart, Y. B., & Grant, M. C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, *27*(1), 237–277.

Liu, S., Chen, S., Chen, Y., Guan, Z., Yin, D., & Chen, F. (2011). In vitro induced tetraploid of *Dendranthema nankingense* (Nakai) Tzvel. shows an improved level of abiotic stress tolerance. *Scientia Horticulturae*, *127*(3), 411–419.

Liu, W., Tai, H., Li, S., Gao, W., Zhao, M., Xie, C., & Li, W.-X. (2014). bHLH122 is important for drought and osmotic stress resistance in Arabidopsis and in the repression of ABA catabolism. *The New Phytologist*, *201*(4), 1192–1204.

Liu, W., Zhao, Y., You, J., Qi, D., Zhou, Y., Chen, J., & Song, Z. (2016). Morphological and genetic variation along a north-to-south transect in *Sipa purpurea*, a dominant grass on the qinghai-tibetan plateau: implications for response to climate change. *PloS One*, *11*(8), e0161972.

Li, Y., Huang, Y., Bergelson, J., Nordborg, M., & Borevitz, J. O. (2010). Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(49), 21199–21204.

Li, Z., Baniaga, A. E., Sessa, E. B., Scascitelli, M., Graham, S. W., Rieseberg, L. H., & Barker, M. S. (2015). Early genome duplications in conifers and other seed plants. *Science Advances*, *1*(10), e1501084.

Lloret, F., Peñuelas, J., & Estiarte, M. (2004). Experimental evidence of reduced diversity of seedlings due to climate modification in a Mediterranean-type community. In *Global Change Biology* (Vol. 10, Issue 2, pp. 248–258). https://doi.org/10.1111/j.1365-2486.2004.00725.x

Loha, A., Tigabu, M., & Teketay, D. (2008). Variability in seed- and seedling-related traits of *Millettia ferruginea*, a potential agroforestry species. *New Forests*, *36*(1), 67–78.

López, A. S., López, D. R., Caballé, G., Siffredi, G. L., & Marchelli, P. (2020). Local adaptation along a sharp rainfall gradient occurs in a native Patagonian grass, *Festuca pallescens*, regardless of extensive gene flow. *Environmental and Experimental Botany*, *171*, 103933.

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, *23*(9), 2178–2192.

Lotterhos, K. E., & Whitlock, M. C. (2015). The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, *24*(5), 1031–1046.

Lourkisti, R., Froelicher, Y., Herbette, S., Morillon, R., Tomi, F., Gibernau, M., Giannettini, J., Berti, L., & Santini, J. (2020). Triploid citrus genotypes have a better tolerance to natural chilling conditions of photosynthetic capacities and specific leaf volatile organic compounds. *Frontiers in Plant Science*, *11*, 330.

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

Lovell, J. T., MacQueen, A. H., Mamidi, S., Bonnette, J., Jenkins, J., Napier, J. D., Sreedasyam, A., Healey, A., Session, A., Shu, S., Barry, K., Bonos, S., Boston, L., Daum, C., Deshpande, S., Ewing, A., Grabowski, P. P., Haque, T., Harrison, M., … Schmutz, J. (2021). Genomic mechanisms of climate adaptation in polyploid bioenergy switchgrass. *Nature*, *590*(7846), 438–444.

Lowe, A. J., Boshier, D., Ward, M., Bacles, C. F., & Navarro, C. (2005). Genetic resource impacts of habitat loss and degradation; reconciling empirical evidence and predicted theory for Neotropical trees. *Heredit 95: 255-273*. Chapman and Hall.

Lowry, D. B., Purmal, C. T., & Juenger, T. E. (2013). A population genetic transect of *Panicum hallii* (Poaceae). *American Journal of Botany*, *100*(3), 592–601.

Lu, M., Krutovsky, K. V., & Loopstra, C. A. (2019). Predicting adaptive genetic variation of loblolly pine (*Pinus taeda* L.) populations under projected future climates based on multivariate models. *The Journal of Heredity*, *110*(7), 857–865.

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*(12), 981–994.

Lunt, I. D., & Morgan, J. W. (2002). The role of fire regimes in temperate lowland grasslands of southeastern Australia. *: The Fire Regimes and Biodiversity of a …*. https://books.google.com/books?hl=en&lr=&id=4f8anIoS33MC&oi=fnd&pg=PA177&dq=The+role

+of+fire+regimes+in+temperate+lowland+grasslands+of+southeastern+Australia&ots=kHerGYsYmn&sig=XVb0MRKxOq3Jzy_okk9eF0v4bfU

Lynch, M., Conery, J., & Burger, R. (1995). Mutation Accumulation and the Extinction of Small Populations. *The American Naturalist*, *146*(4), 489–518.

Maere, S., De Bodt, S., Raes, J., Casneuf, T., Van Montagu, M., Kuiper, M., & Van de Peer, Y. (2005). Modeling gene and genome duplications in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(15), 5454–5459.

Maherali, H., Walden, A. E., & Husband, B. C. (2009). Genome duplication and the evolution of physiological responses to water stress. *The New Phytologist*, *184*(3), 721–731.

Mahmood, T., Iqbal, M. S., Li, H., Nazir, M. F., Khalid, S., Sarfraz, Z., Hu, D., Baojun, C., Geng, X., Tajo, S. M., Dev, W., Iqbal, Z., Zhao, P., Hu, G., & Du, X. (2022). Differential seedling growth and tolerance indices reflect drought tolerance in cotton. *BMC Plant Biology*, *22*(1), 331.

Mahony, C. R., MacLachlan, I. R., Lind, B. M., Yoder, J. B., Wang, T., & Aitken, S. N. (2019). Evaluating genomic data for management of local adaptation in a changing climate: A lodgepole pine case study. https://doi.org/10.1101/568725

Ma, J., & Bennetzen, J. L. (2004). Rapid recent growth and divergence of rice nuclear genomes. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(34), 12404–12410.

Malyshev, A. V., Arfin Khan, M. A. S., Beierkuhnlein, C., Steinbauer, M. J., Henry, H. A. L., Jentsch, A., Dengler, J., Willner, E., & Kreyling, J. (2016). Plant responses to climatic extremes: within-species variation equals among-species variation. *Global Change Biology*, *22*(1), 449–464.

Mamo, N., Mihretu, M., Fekadu, M., Tigabu, M., & Teketay, D. (2006). Variation in seed and germination characteristics among *Juniperus procera* populations in Ethiopia. *Forest Ecology and Management*, *225*(1), 320–327.

Mandáková, T., Joly, S., Krzywinski, M., Mummenhoff, K., & Lysak, M. A. (2010). Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. In *The Plant Cell* (Vol. 22, Issue 7, pp. 2277–2290). https://doi.org/10.1105/tpc.110.074526

Manel, S., Joost, S., Epperson, B. K., Holderegger, R., Storfer, A., Rosenberg, M. S., Scribner, K. T., Bonin, A., & Fortin, M.-J. (2010). Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, *19*(17), 3760–3772.

Manel, S., Perrier, C., Pratlong, M., Abi-Rached, L., Paganini, J., Pontarotti, P., & Aurelle, D. (2016). Genomic resources and their influence on the detection of the signal of positive selection in genome scans. *Molecular Ecology*, *25*(1), 170–184.

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., & Zdobnov, E. M. (2021). BUSCO Update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution*, *38*(10), 4647–4654.

Mao, Y. (2019). GenoDup Pipeline: a tool to detect genome duplication using the dS-based method. *PeerJ*, *7*, e6303.

Marçais, G., & Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* , *27*(6), 764–770.

Marques, I., Shiposha, V., López-Alvarez, D., Manzaneda, A. J., Hernandez, P., Olonova, M., & Catalán, P. (2017). Environmental isolation explains Iberian genetic diversity in the highly homozygous model grass Brachypodium distachyon. *BMC Evolutionary Biology*, *17*(1), 139.

Martienssen, R. A., & Colot, V. (2001). DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science*, *293*(5532), 1070–1074.

Masterson, J. (1994). Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science*, *264*(5157), 421–424.

Mayrose, I., Zhan, S. H., Rothfels, C. J., Magnuson-Ford, K., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2011). Recently formed polyploid plants diversify at lower rates. *Science*, *333*(6047), 1257.

McCarthy, E. W., Chase, M. W., Knapp, S., Litt, A., Leitch, A. R., & Le Comber, S. C. (2016). Transgressive phenotypes and generalist pollination in the floral evolution of *Nicotiana* polyploids. *Nature Plants*, *2*, 16119.

McCartney-Melstad, E., & Shaffer, H. B. (2015). Amphibian molecular ecology and how it has informed conservation. *Molecular Ecology*, *24*(20), 5084–5109.

McKay, J. K., Bishop, J. G., Lin, J. Z., Richards, J. H., Sala, A., & Mitchell-Olds, T. (2001). Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proceedings. Biological Sciences / The Royal Society*, *268*(1477), 1715–1721.

McKay, J. K., & Latta, R. G. (2002). Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*, *17*(6), 285–291.

McKinney, G. J., Larson, W. A., Seeb, L. W., & Seeb, J. E. (2017). RADseq provides unprecedented insights into molecular ecology and evolutionary genetics. *Molecular Ecology Resources*, *17*(3), 356–361.

McKinney, G. J., Waples, R. K., Seeb, L. W., & Seeb, J. E. (2017). Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Molecular Ecology Resources*, *17*(4), 656–669.

Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, *21*(12), 2839–2846.

Meirmans, P. G., Liu, S., & van Tienderen, P. H. (2018). The analysis of polyploid genetic data. *The Journal of Heredity*, *109*(3), 283–296.

Meirmans, P. G., & Van Tienderen, P. H. (2013). The effects of inheritance in tetraploids on genetic diversity and population divergence. *Heredity*, *110*(2), 131–137.

Méndez-Toribio, M., Ibarra-Manríquez, G., Paz, H., & Lebrija-Trejos, E. (2020). Atmospheric and soil drought risks combined shape community assembly of trees in a tropical dry forest. *The Journal of Ecology*, *108*(4), 1347–1357.

Merilä, J., & Crnokrak, P. (2001). Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, *14*(6), 892–903.

Merilä, J., & Sheldon, B. C. (1999). Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. *Heredity*, *83 (Pt 2)*, 103–109.

Meyer, S. E., Leger, E. A., Eldon, D. R., & Coleman, C. E. (2016). Strong genetic differentiation in the invasive annual grass *Bromus tectorum* across the Mojave–Great Basin ecological transition zone. *Biological Invasions*, *18*(6), 1611–1628.

Miller, J. R., Wood, B. P., & Hamilton, M. B. (2008). F(ST) and Q(ST) under neutrality. *Genetics*, *180*(2), 1023–1037.

Mitchell, M. (2002). Native Grasses: An Identification Handbook for Temperate Australia. Landlinks Press.

Mitros, T., Session, A. M., James, B. T., Wu, G. A., Belaffif, M. B., Clark, L. V., Shu, S., Dong, H., Barling, A., Holmes, J. R., Mattick, J. E., Bredeson, J. V., Liu, S., Farrar, K., Głowacka, K., Jeżowski, S., Barry, K., Chae, W. B., Juvik, J. A., … Rokhsar, D. S. (2020). Genome biology of the paleotetraploid perennial biomass crop *Miscanthus*. *Nature Communications*, *11*(1), 5442.

Moniz de Sá, M., & Drouin, G. (1996). Phylogeny and substitution rates of angiosperm actin genes. *Molecular Biology and Evolution*, *13*(9), 1198–1212.

Montwé, D., Spiecker, H., & Hamann, A. (2015). Five decades of growth in a genetic field trial of Douglas-fir reveal trade-offs between productivity and drought tolerance. *Tree Genetics & Genomes*, *11*(2), 29.

Moore, C. (1957). A study of some soil properties in relation to the invasion of native pastures by *Bothriochloa ambigua* S.T. Blake. In *Australian Journal of Botany* (Vol. 5, Issue 1, p. 44). https://doi.org/10.1071/bt9570044

Morin, P. A., Luikart, G., Wayne, R. K., & the SNP workshop group. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, *19*(4), 208–216.

Morrissey, M. B., Walling, C. A., Wilson, A. J., Pemberton, J. M., Clutton-Brock, T. H., & Kruuk, L. E. B. (2012). Genetic analysis of life-history constraint and evolution in a wild ungulate population. *The American Naturalist*, *179*(4), E97–E114.

Mueller, J. M., & Hellmann, J. J. (2008). An assessment of invasion risk from assisted migration. *Conservation Biology: The Journal of the Society for Conservation Biology*, *22*(3), 562–567.

Muktar, M. S., Teshome, A., Hanson, J., Negawo, A. T., Habte, E., Domelevo Entfellner, J.-B., Lee, K.-W., & Jones, C. S. (2019). Genotyping by sequencing provides new insights into the diversity of Napier grass (*Cenchrus purpureus*) and reveals variation in genome-wide LD patterns between collections. *Scientific Reports*, *9*(1), 6936.

Nagamitsu, T., Shimada, K.-I., & Kanazashi, A. (2014). A reciprocal transplant trial suggests a disadvantage of northward seed transfer in survival and growth of Japanese red pine (*Pinus densiflora*) trees. *Tree Genetics & Genomes*, *11*(1), 813.

Narum, S. R., & Hess, J. E. (2011). Comparison of F(ST) outlier tests for SNP loci under selection. *Molecular Ecology Resources*, *11 Suppl 1*, 184–194.

Nguyen-Queyrens, A., & Bouchet-Lannat, F. (2003). Osmotic adjustment in three-year-old seedlings of five provenances of maritime pine (*Pinus pinaster*) in response to drought. *Tree Physiology*, *23*(6), 397–404.

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., & van Kleunen, M. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, *15*(12), 684–692.

Nik, M. M., Babaeian, M., Tavassoli, A., & Others. (2011). Effect of seed size and genotype on germination characteristic and seed nutrient content of wheat. *Scientific Research and Essays*, *6*(9), 2019–2025.

Nizam, I. (2011). Effects of salinity stress on water uptake, germination and early seedling growth of perennial ryegrass. *African Journal of Biotechnology*, *10*(51), 10418–10424.

Nunez, S., Arets, E., Alkemade, R., Verwer, C., & Leemans, R. (2019). Assessing the impacts of climate change on biodiversity: is below 2 °C enough? *Climatic Change*, *154*(3), 351–365.

Ohno, S. (1970). Evolution by Gene Duplication. Springer Science & Business Media.

Oliver, T. H., Smithers, R. J., Bailey, S., Walmsley, C. A., & Watts, K. (2012). A decision framework for considering climate change adaptation in biodiversity conservation planning. In *Journal of Applied Ecology* (Vol. 49, Issue 6, pp. 1247–1255). https://doi.org/10.1111/1365-2664.12003

O'Neill, G. A., Stoehr, M., & Jaquish, B. (2014). Quantifying safe seed transfer distance and impacts of tree breeding on adaptation. *Forest Ecology and Management*, *328*, 122–130.

Orsini, L., Mergeay, J., Vanoverbeke, J., & De Meester, L. (2013). The role of selection in driving landscape genomic structure of the waterflea *Daphnia magna*. *Molecular Ecology*, *22*(3), 583–601.

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics*, *34*, 401–437.

Oyundelger, K., Ritz, C. M., Munkhzul, O., Lang, B., Ahlborn, J., Oyuntsetseg, B., Römermann, C., & Wesche, K. (2020). Climate and land use affect genetic structure of *Stipa glareosa* P. A. Smirn. in Mongolia. In *Flora* (Vol. 266, p. 151572). https://doi.org/10.1016/j.flora.2020.151572

Pacifici, M., Foden, W. B., Visconti, P., Watson, J. E. M., Butchart, S. H. M., Kovacs, K. M., Scheffers, B. R., Hole, D. G., Martin, T. G., Resit Akçakaya, H., Corlett, R. T., Huntley, B., Bickford, D., Carr, J. A., Hoffmann, A. A., Midgley, G. F., Pearce-Kelly, P., Pearson, R. G., Williams, S. E., … Rondinini, C. (2015). Assessing species vulnerability to climate change. In *Nature Climate Change* (Vol. 5, Issue 3, pp. 215–224). https://doi.org/10.1038/nclimate2448

Panchy, N., Lehti-Shiu, M., & Shiu, S.-H. (2016). Evolution of gene duplication in plants. *Plant Physiology*, *171*(4), 2294–2316.

Paradis, E., & Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* , *35*(3), 526–528.

Parkhurst, M. M. J., Doust, A., Mauro-Herrera, M., Steets, J. A., & Byrnes, J. M. (2011). Spatial genetic structure of the tallgrass prairie grass *dichanthelium oligosanthes* (scribner's panicum). *Oklahoma Native Plant Record*, *11*(1). https://doi.org/10.22488/okstate.17.100083

Paterson, A. H., Chapman, B. A., Kissinger, J. C., Bowers, J. E., Feltus, F. A., & Estill, J. C. (2006). Many gene and domain families have convergent fates following independent whole-genome duplication events in *Arabidopsis*, *Oryza*, *Saccharomyces* and *Tetraodon*. *Trends in Genetics: TIG*, *22*(11), 597–602.

Pearman, W. S., Urban, L., & Alexander, A. (2022). Commonly used Hardy-Weinberg equilibrium filtering schemes impact population structure inferences using RADseq data. *Molecular Ecology Resources*, *22*(7), 2599–2613.

Perrier, C., Lozano del Campo, A., Szulkin, M., Demeyrier, V., Gregoire, A., & Charmantier, A. (2018). Great tits and the city: Distribution of genomic diversity and gene-environment associations along an urbanization gradient. *Evolutionary Applications*, *11*(5), 593–613.

Pina-Martins, F., Baptista, J., & Pappas, G., Jr. (2019). New insights into adaptation and population structure of cork oak using genotyping by sequencing. *Global Change Biology*. https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.14497

Pirnajmedin, F., Majidi, M. M., Saeidi, G., Gheysari, M., Volaire, F., Barre, P., Osivand, A. H., & Sarfaraz, D. (2017). Persistence, recovery and root traits of tall fescue genotypes with different

flowering date under prolonged water stress. *Euphytica/ Netherlands Journal of Plant Breeding*, *213*(12), 269.

Pitman, A. J., Narisma, G. T., & McAneney, J. (2007). The impact of climate change on the risk of forest and grassland fires in Australia. *Climatic Change*, *84*(3), 383–401.

Plant Protein Database. Retrieved November 10, 2020, from http://www.hrt.msu.edu/uploads/535/78637/alluniRefprexp070416.gz

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PloS One*, *7*(2), e32253.

Poplin, R., Ruano-Rubio, V., DePristo, M. A., & Fennell, T. J. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*. https://www.biorxiv.org/content/10.1101/201178v2.abstract

Power, S. A., Barnett, K. L., Ochoa-Hueso, R., Facey, S. L., Gibson-Forty, E. V. J., Hartley, S. E., Nielsen, U. N., Tissue, D. T., & Johnson, S. N. (2016). DRI-Grass: A new experimental platform for addressing grassland ecosystem responses to future precipitation scenarios in south-east australia. *Frontiers in Plant Science*, *7*, 1373.

Pramova, E., Locatelli, B., Djoudi, H., Lavorel, S., Colloff, M. J., & Martius, C. (2019). Adapting land restoration to a changing climate: Embracing the knowns and unknowns. CIFOR.

Prober, Gosper, & Gilfedder. (2017). Temperate eucalypt woodlands. *Vegetation History and Archaeobotany*. https://books.google.com/books?hl=en&lr=&id=Dk3ODgAAQBAJ&oi=fnd&pg=PA410&dq=Prober+S+M+Gosper+C+R+Gilfedder+L+(2017)+Temperate+Eucalypt+Woodlands+In+Australian+Vegetation+3rd+edn+(ed+D+Keith)+Cambridge+University+Press+Cambridge+UK&ots=JVzj0K2mF0&sig=y2LaIG0bm2J5Yy6rMYnJlfnhnbc

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. In *Frontiers in Ecology and Evolution* (Vol. 3). https://doi.org/10.3389/fevo.2015.00065

Prober, S. M., & Thiele, K. R. (2005). Restoring Australia's temperate grasslands and grassy woodlands: integrating function and diversity. *Ecological Management & Restoration*, *6*(1), 16–27.

Prober, S. M., Thiele, K. R., & Lunt, I. D. (2002). Identifying ecological barriers to restoration in temperate grassy woodlands: soil changes associated with different degradation states. *Australian Journal of Botany*, *50*(6), 699–712.

Prunier, R., Akman, M., Kremer, C. T., Aitken, N., Chuah, A., Borevitz, J., & Holsinger, K. E. (2017). Isolation by distance and isolation by environment contribute to population differentiation in Protea repens (Proteaceae L.), a widespread South African species. *American Journal of Botany*, *104*(5), 674–684.

Pujol, B., Wilson, A. J., Ross, R. I. C., & Pannell, J. R. (2008). Are QST–FST comparisons for natural populations meaningful? *Molecular Ecology*, *17*(22), 4782–4785.

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., Pang, N., Forslund, K., Ceric, G., Clements, J., Heger, A., Holm, L., Sonnhammer, E. L. L., Eddy, S. R., Bateman, A., & Finn, R. D. (2012). The Pfam protein families database. *Nucleic Acids Research*, *40*(Database issue), D290–D301.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Putnam, N. H., O'Connell, B. L., Stites, J. C., Rice, B. J., Blanchette, M., Calef, R., Troll, C. J., Fields, A., Hartley, P. D., Sugnet, C. W., Haussler, D., Rokhsar, D. S., & Green, R. E. (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Research*, *26*(3), 342–350.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., & Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Research*, *33*(Web Server issue), W116–W120.

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904.

Ramsey, J. (2011). Polyploidy and ecological adaptation in wild yarrow. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(17), 7096–7101.

Ramsey, J., & Ramsey, T. S. (2014). Ecological studies of polyploidy in the 100 years following its discovery. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*(1648). https://doi.org/10.1098/rstb.2013.0352

Ramsey, J., & Schemske, D. W. (2003). *Neopolyploidy in Flowering Plants*. https://doi.org/10.1146/annurev.ecolsys.33.010802.150437

Rauschkolb, R., Li, Z., Godefroid, S., Dixon, L., Durka, W., Májeková, M., Bossdorf, O., Ensslin, A., & Scheepens, J. F. (2022). Evolution of plant drought strategies and herbivore tolerance after two decades of climate change. *The New Phytologist*, *235*(2), 773–785.

Razzaque, S., & Juenger, T. E. (2022). The ecology and quantitative genetics of seed and seedling traits in upland and lowland ecotypes of a perennial grass. *Evolution Letters*, *6*(6), 460–473.

Reed, D. H. (2003). Inbreeding and extinction: effects of rate of inbreeding. *Conservation Genetics* , *4*(3), 405–410.

Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., & Hamilton, D. A., Jr. (1999). Genetic responses to climate inpinus contorta: Niche breadth, climate change, and reforestation. *Ecological Monographs*, *69*(3), 375–407.

Rellstab, C., Dauphin, B., & Exposito-Alonso, M. (2021). Prospects and limitations of genomic offset in conservation management. *Evolutionary Applications*, *14*(5), 1202–1212.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, *24*(17), 4348–4370.

Renny-Byfield, S., Rodgers-Melnick, E., & Ross-Ibarra, J. (2017). Gene fractionation and function in the ancient subgenomes of maize. *Molecular Biology and Evolution*, *34*(8), 1825–1832.

Reside, A. E., Butt, N., & Adams, V. M. (2018). Adapting systematic conservation planning for climate change. *Biodiversity and Conservation*, *27*(1), 1–29.

Rey Benayas, J. M., Newton, A. C., Diaz, A., & Bullock, J. M. (2009). Enhancement of biodiversity and ecosystem services by ecological restoration: a meta-analysis. *Science*, *325*(5944), 1121–1124.

Reynolds, R. G., Puente-Rolón, A. R., Platenberg, R., Tyler, R. K., Tolson, P. J., & Revell, L. J. (2015). Large divergence and low diversity suggest genetically informed conservation strategies for the endangered Virgin Islands Boa (*Chilabothrus monensis*). *Global Ecology and Conservation*, *3*, 487–502.

Ricciardi, A., & Simberloff, D. (2009). Assisted colonization is not a viable conservation strategy. *Trends in Ecology & Evolution*, *24*(5), 248–253.

Rice, A., Šmarda, P., Novosolov, M., Drori, M., Glick, L., Sabath, N., Meiri, S., Belmaker, J., & Mayrose, I. (2019). The global biogeography of polyploid plants. In *Nature Ecology & Evolution* (Vol. 3, Issue 2, pp. 265–273). https://doi.org/10.1038/s41559-018-0787-9

Rieseberg, L. H., Raymond, O., Rosenthal, D. M., Lai, Z., Livingstone, K., Nakazato, T., Durphy, J. L., Schwarzbach, A. E., Donovan, L. A., & Lexer, C. (2003). Major ecological transitions in wild sunflowers facilitated by hybridization. *Science*, *301*(5637), 1211–1216.

Robinson, J. T., Turner, D., Durand, N. C., Thorvaldsdóttir, H., Mesirov, J. P., & Aiden, E. L. (2018). Juicebox.js provides a cloud-based visualization system for Hi-C data. *Cell Systems*, *6*(2), 256–258.e1.

Rody, H. V. S., Baute, G. J., Rieseberg, L. H., & Oliveira, L. O. (2017). Both mechanism and age of duplications contribute to biased gene retention patterns in plants. *BMC Genomics*, *18*(1), 46.

Rognli, O. A., Nilsson, N. O., & Nurminiemi, M. (2000). Effects of distance and pollen competition on gene flow in the wind-pollinated grass *Festuca pratensis* Huds. *Heredity*, *85*(Pt 6), 550–560.

Rossetto, & Rymer. (2013). Applications of molecular markers in plant conservation. *Molecular Markers in Plants*. https://doi.org/10.1002/9781118473023#page=96

Ross-Ibarra, J., Morrell, P. L., & Gaut, B. S. (2007). Plant domestication, a unique opportunity to identify the genetic basis of adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, *104 Suppl 1*, 8641–8648.

Roybal, C. M., & Butterfield, B. J. (2018). Functional trait heritability and local climatic adaptation among grasses: a meta-analysis. *Plant Ecology*, *219*(4), 369–379.

Rudmann-Maurer, K., Weyand, A., Fischer, M., & Stöcklin, J. (2007). Microsatellite diversity of the agriculturally important alpine grass *Poa alpina* in relation to land use and natural environment. *Annals of Botany*, *100*(6), 1249–1258.

Sambatti, J. B. M., & Rice, K. J. (2006). Local adaptation, patterns of selection, and gene flow in the Californian serpentine sunflower (*Helianthus exilis*). *Evolution; International Journal of Organic Evolution*, *60*(4), 696–710.

Sankoff, D., Zheng, C., & Zhu, Q. (2010). The collapse of gene complement following whole genome duplication. *BMC Genomics*, *11*, 313.

Saremi, N. F., Supple, M. A., Byrne, A., Cahill, J. A., Coutinho, L. L., Dalén, L., Figueiró, H. V., Johnson, W. E., Milne, H. J., O'Brien, S. J., O'Connell, B., Onorato, D. P., Riley, S. P. D., Sikich, J. A., Stahler, D. R., Villela, P. M. S., Vollmers, C., Wayne, R. K., Eizirik, E., … Shapiro, B. (2019). Puma genomes from North and South America provide insights into the genomic consequences of inbreeding. *Nature Communications*, *10*(1), 4769.

Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews. Genetics*, *14*(11), 807–820.

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 595–619.

Scarpino, S. V., Levin, D. A., & Meyers, L. A. (2014). Polyploid formation shapes flowering plant diversity. *The American Naturalist*, *184*(4), 456–465.

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*(6313). https://doi.org/10.1126/science.aaf7671

Schemske, D. W., Husband, B. C., Ruckelshaus, M. H., Goodwillie, C., Parker, I. M., & Bishop, J. G. (1994). Evaluating approaches to the conservation of rare and endangered plants. *Ecology*, *75*(3), 585–606.

Schiffers, K., Bourne, E. C., Lavergne, S., Thuiller, W., & Travis, J. M. J. (2013). Limited evolutionary rescue of locally adapted populations facing climate change. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *368*(1610), 20120083.

Schlaepfer, D. R., Bradford, J. B., Lauenroth, W. K., Munson, S. M., Tietjen, B., Hall, S. A., Wilson, S. D., Duniway, M. C., Jia, G., Pyke, D. A., Lkhagva, A., & Jamiyansharav, K. (2017). Climate change reduces extent of temperate drylands and intensifies drought in deep soils. *Nature Communications*, *8*, 14196.

Schmickl, R., Marburger, S., Bray, S., & Yant, L. (2017). Hybrids and horizontal transfer: introgression allows adaptive allele discovery. *Journal of Experimental Botany*, *68*(20), 5453–5470.

Schnable, J. C., Springer, N. M., & Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(10), 4069–4074.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., Liang, C., Zhang, J., Fulton, L., Graves, T. A., Minx, P., Reily, A. D., Courtney, L., Kruchowski, S. S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., … Wilson, R. K. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science*, *326*(5956), 1112–1115.

Schwartz, M. W., Hellmann, J. J., & McLachlan, J. S. (2009). The precautionary principle in managed relocation is misguided advice *Trends in Ecology & Evolution*, *24*(9), 474.

Sémon, M., & Wolfe, K. H. (2007). Consequences of genome duplication. *Current Opinion in Genetics & Development*, *17*(6), 505–512.

Semple, W. S., Koen, T. B., & Waterhouse, D. (1997). Consequences of some one-off events and exclosure on a red grass (*Bothriochloa macra*)-wallaby grass (*Danthonia eriantha*) pasture in the central west of NSW. *Rangeland Journal*, *19*(2), 206–215.

Seoighe, C., & Gehring, C. (2004). Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome. *Trends in Genetics: TIG*, *20*(10), 461–464.

Seppey, M., Manni, M., & Zdobnov, E. M. (2019). BUSCO: Assessing genome assembly and annotation completeness. *Methods in Molecular Biology , 1962*, 227–245.

Session, A. M., Uno, Y., Kwon, T., Chapman, J. A., Toyoda, A., Takahashi, S., Fukui, A., Hikosaka, A., Suzuki, A., Kondo, M., van Heeringen, S. J., Quigley, I., Heinz, S., Ogino, H., Ochi, H., Hellsten, U., Lyons, J. B., Simakov, O., Putnam, N., … Rokhsar, D. S. (2016). Genome evolution in the allotetraploid frog *Xenopus laevis*. *Nature*, *538*(7625), 336–343.

Sexton, J. P., Hangartner, S. B., & Hoffmann, A. A. (2014). Genetic isolation by environment or distance: which pattern of gene flow is most common? *Evolution; International Journal of Organic Evolution*, *68*(1), 1–15.

Sexton, J. P., Strauss, S. Y., & Rice, K. J. (2011). Gene flow increases fitness at the warm edge of a species' range. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(28), 11704–11709.

Sgrò, C. M., Lowe, A. J., & Hoffmann, A. A. (2011). Building evolutionary resilience for conserving biodiversity under climate change. *Evolutionary Applications*, *4*(2), 326–337.

Shafer, A. B. A., Peart, C. R., & Tusso, S. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution / British Ecological Society*. https://doi.org/10.1111/2041-210X.12700

Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology. *Ecology Letters*, *16*(7), 940–950.

Shantz, H. L. (1954). The Place of Grasslands in the Earth's Cover. *Ecology*, *35*(2), 143–145.

Simon, B. K., & Alfonso, Y. (2011). *AusGrass2. Brisbane, Queensland, Australia*.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, *236*(4803), 787–792.

Slatkin, M. (1993). Isolation by distance in equilibrium and non-equilibrium populations. *Evolution; International Journal of Organic Evolution*, *47*(1), 264–279.

Sloat, L. L., Gerber, J. S., Samberg, L. H., Smith, W. K., Herrero, M., Ferreira, L. G., Godde, C. M., & West, P. C. (2018). Increasing importance of precipitation variability on global livestock grazing lands. *Nature Climate Change*, *8*(3), 214–218.

Smit, A. F. A., & Hubley, R. (2008). *RepeatModeler Open-1.0 http://www. repeatmasker. org*. RepeatModeler.

Smit, A. F. A., Hubley, R., & Green, P. (2015). *RepeatMasker Open-4.0. 2013--2015*.

Soltis, D. E., Buggs, R. J. A., Doyle, J. J., & Soltis, P. S. (2010). What we still don't know about polyploidy. *Taxon*, *59*(5), 1387–1403.

Soltis, D. E., Visger, C. J., Marchant, D. B., & Soltis, P. S. (2016). Polyploidy: Pitfalls and paths to a paradigm. *American Journal of Botany*, *103*(7), 1146–1166.

Soltis, D. E., Visger, C. J., & Soltis, P. S. (2014). The polyploidy revolution then…and now: Stebbins revisited. *American Journal of Botany*, *101*(7), 1057–1078.

Soltis, P. S., & Soltis, D. E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology*, *30*, 159–165.

Song, B.-H., Clauss, M. J., Pepper, A., & Mitchell-Olds, T. (2006). Geographic patterns of microsatellite variation in *Boechera stricta*, a close relative of Arabidopsis. *Molecular Ecology*, *15*(2), 357–369.

Sork, V. L., Aitken, S. N., Dyer, R. J., Eckert, A. J., Legendre, P., & Neale, D. B. (2013). Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, *9*(4), 901–911.

Spitze, K. (1993). Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*, *135*(2), 367–374.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* , *30*(9), 1312–1313.

Stange, M., Barrett, R. D. H., & Hendry, A. P. (2021). The importance of genomic variation for biodiversity, ecosystems and people. *Nature Reviews. Genetics*, *22*(2), 89–105.

Stanke, M., Schöffmann, O., Morgenstern, B., & Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*, *7*, 62.

Stanley, T. D., & Ross, E. M. (1983). Flora of south-eastern Queensland. Queensland Dept. of Primary Industries.

Steane, D. A., Mclean, E. H., Potts, B. M., Prober, S. M., Stock, W. D., Stylianou, V. M., Vaillancourt, R. E., & Byrne, M. (2017). Evidence for adaptation and acclimation in a widespread eucalypt of semi-arid Australia. *Biological Journal of the Linnean Society. Linnean Society of London*, *121*(3), 484–500.

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*(10), 2500–2513.

Stebbins, G. L. (1971). Processes of organic evolution. *Second Edition*. Prentice-Hall.

Stebbins, G. L., & Ledyard Stebbins, G. (1985). Polyploidy, hybridization, and the invasion of new habitats. In *Annals of the Missouri Botanical Garden* (Vol. 72, Issue 4, p. 824). https://doi.org/10.2307/2399224

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., de Vries, W., de Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., & Sörlin, S. (2015). Sustainability. Planetary boundaries: guiding human development on a changing planet. *Science*, *347*(6223), 1259855.

Steinbiss, S., Willhoeft, U., Gremme, G., & Kurtz, S. (2009). Fine-grained annotation and classification of de novo predicted LTR retrotransposons. *Nucleic Acids Research*, *37*(21), 7002–7013.

Stoffel, M. A., Esser, M., Kardos, M., Humble, E., Nichols, H., David, P., & Hoffman, J. I. (2016). inbreedR: an R package for the analysis of inbreeding based on genetic markers. In *Methods in Ecology and Evolution* (Vol. 7, Issue 11, pp. 1331–1339). https://doi.org/10.1111/2041-210x.12588

Storer, J., Hubley, R., Rosen, J., Wheeler, T. J., & Smit, A. F. (2021). The Dfam community resource of transposable element families, sequence models, and genome annotations. *Mobile DNA*, *12*(1), 2.

Stroud, H., Greenberg, M. V. C., Feng, S., Bernatavichute, Y. V., & Jacobsen, S. E. (2013). Comprehensive analysis of silencing mutants reveals complex regulation of the Arabidopsis methylome. *Cell*, *152*(1-2), 352–364.

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*(1), vey016.

Sumadijaya, A. (2015). Morphology, molecular phylogeny and genome content of *Bothriochloa* focusing on australian taxa [Virginia Tech]. https://vtechworks.lib.vt.edu/handle/10919/73657

Supple, M. A., & Shapiro, B. (2018). Conservation of biodiversity in the genomics era. *Genome Biology*, *19*(1), 131.

Šurinová, M., Hadincová, V., Vandvik, V., & Münzbergová, Z. (2019). Temperature and precipitation, but not geographic distance, explain genetic relatedness among populations in the perennial grass *Festuca rubra*. *Journal of Plant Ecology*, *12*(4), 730–741.

Sweigart, A. L., & Willis, J. H. (2003). Patterns of nucleotide diversity in two species of *Mimulus* are affected by mating system and asymmetric introgression. *Evolution; International Journal of Organic Evolution*, *57*(11), 2490–2506.

Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–577.

Tasdighian, S., Van Bel, M., Li, Z., Van de Peer, Y., Carretero-Paulet, L., & Maere, S. (2017). Reciprocally retained genes in the angiosperm lineage show the hallmarks of dosage balance sensitivity. *The Plant Cell*, *29*(11), 2766–2785.

Tayalé, A., & Parisod, C. (2013). Natural pathways to polyploidy in plants and consequences for genome reorganization. *Cytogenetic and Genome Research*, *140*(2-4), 79–96.

Thammina, C. S., Amundsen, K., Bushman, S. B., Kramer, M., & Warnke, S. E. (2018). Genetic diversity of *Danthonia spicata* (L.) Beauv. based on genomic simple sequence repeat markers. *Genetic Resources and Crop Evolution*, *65*(4), 1059–1070.

Thomas, B. C., Pedersen, B., & Freeling, M. (2006). Following tetraploidy in an Arabidopsis ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. *Genome Research*, *16*(7), 934–946.

Timbal, B. (2009). The continuing decline in south-east Australian rainfall—Update to May 2009. *CAWCR Research Letters*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.307.7389&rep=rep1&type=pdf#page=4

Tobin, M. F., Lopez, O. R., & Kursar, T. A. (1999). Responses of tropical understory plants to a severe drought: Tolerance and avoidance of water Stress1. *Biotropica*, *31*(4), 570–578.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G., & Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, *9*(7), 892–908.

Torkamaneh, D., Laroche, J., & Belzile, F. (2016). Genome-Wide SNP Calling from Genotyping by Sequencing (GBS) Data: A Comparison of Seven Pipelines and Two Sequencing Technologies. *PloS One*, *11*(8), e0161333.

Trabucco, & Zomer. (2018). Global aridity index and potential evapotranspiration (ET0) climate database v2. *CGIAR Consort Spat Inf.* https://classes.engr.oregonstate.edu/cce/spring2019/ce202/Data/global-ai_et0/ai_et0/Global%20AI_PET%20v2%20-%20Readme.pdf.

Tungate, K. D., Burton, M. G., Susko, D. J., Sermons, S. M., & Rufty, T. W. (2006). Altered weed reproduction and maternal effects under low-nitrogen fertility. *Weed Science*, *54*(5), 847–853.

Turcotte, M. M., & Levine, J. M. (2016). Phenotypic plasticity and species coexistence. *Trends in Ecology & Evolution*, *31*(10), 803–813.

UniProt Consortium. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research*, *49*(D1), D480–D489.

Urban, M. C. (2015). Accelerating extinction risk from climate change. *Science*, *348*(6234), 571–573.

Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews. Genetics*, *18*(7), 411–424.

van der Meer, S., & Jacquemyn, H. (2015). Genetic diversity and spatial genetic structure of the grassland perennial *saxifraga granulata* along two river systems. *PloS One*, *10*(6), e0130463.

Van der Putten, W. H. (2012). *Climate Change, Aboveground-Belowground Interactions, and Species' Range Shifts*. https://doi.org/10.1146/annurev-ecolsys-110411-160423

van Dongen, S., & Abreu-Goodger, C. (2012). Using MCL to extract clusters from networks. *Methods in Molecular Biology* , *804*, 281–295.

van Kleunen, M., Fischer, M., & Schmid, B. (2002). Experimental life-history evolution: selection on the allocation to sexual reproduction and its plasticity in a clonal plant. *Evolution; International Journal of Organic Evolution*, *56*(11), 2168–2177.

Varas-Myrik, A., Sepúlveda-Espinoza, F., Fajardo, A., Alarcón, D., Toro-Núñez, Ó., Castro-Nallar, E., & Hasbún, R. (2022). Predicting climate change-related genetic offset for the endangered southern South American conifer *Araucaria araucana*. *Forest Ecology and Management*, *504*, 119856.

Vitt, P., Havens, K., & Hoegh-Guldberg, O. (2009). Assisted migration: part of an integrated conservation strategy [Review of *Assisted migration: part of an integrated conservation strategy*]. *Trends in Ecology & Evolution*, *24*(9), 473–474; author reply 476–477.

von Wettberg, E. J. B., Vance, W., & Rowland, D. L. (2014). The Park Grass Experiment and next-generation approaches: local adaptation of sweet vernal grass revisited [Review of *The Park Grass Experiment and next-generation approaches: local adaptation of sweet vernal grass revisited*]. *Molecular Ecology*, *23*(24), 5931–5933.

Wagner, W. H., Jr. (1970). Biosystematics and evolutionary noise. *Taxon*, *19*(2), 146–151.

Waldvogel, A.-M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., Mock, T., Schmid, K., Schmitt, I., Bataillon, T., Savolainen, O., Bergland, A., Flatt, T., Guillaume, F., & Pfenninger, M. (2020). Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters*, *4*(1), 4–18.

Walters, A. D., & Schwartz, M. K. (2021). Population Genomics for the Management of Wild Vertebrate Populations. In P. A. Hohenlohe & O. P. Rajora (Eds.), *Population Genomics: Wildlife* (pp. 419–436). Springer International Publishing.

Walther, G.-R., Roques, A., Hulme, P. E., Sykes, M. T., Pysek, P., Kühn, I., Zobel, M., Bacher, S., Botta-Dukát, Z., Bugmann, H., Czúcz, B., Dauber, J., Hickler, T., Jarosík, V., Kenis, M., Klotz, S., Minchin, D., Moora, M., Nentwig, W., … Settele, J. (2009). Alien species in a warmer world: risks and opportunities. *Trends in Ecology & Evolution*, *24*(12), 686–693.

Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, *23*(23), 5649–5662.

Wang, W., Shao, A., Xu, X., Fan, S., & Fu, J. (2022). Comparative genomics reveals the molecular mechanism of salt adaptation for zoysiagrasses. *BMC Plant Biology*, *22*(1), 355.

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., Lee, T.-H., Jin, H., Marler, B., Guo, H., Kissinger, J. C., & Paterson, A. H. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, *40*(7), e49.

Warren, R., VanDerWal, J., Price, J., Welbergen, J. A., Atkinson, I., Ramirez-Villegas, J., Osborn, T. J., Jarvis, A., Shoo, L. P., Williams, S. E., & Lowe, J. (2013). Quantifying the benefit of early climate change mitigation in avoiding biodiversity loss. *Nature Climate Change*, *3*(7), 678–682.

Watson, L. (1992). The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry,

cytology, classification, pathogens, world and local distribution, and references. *Http://delta-Intkey. Com*. https://ci.nii.ac.jp/naid/10025924226/

Webber, B. L., Scott, J. K., & Didham, R. K. (2011). Translocation or bust! A new acclimatization agenda for the 21st century? *Trends in Ecology & Evolution*, *26*(10), 495–496

Weber, M., Hellmann, I., Stadler, M. B., Ramos, L., Pääbo, S., Rebhan, M., & Schübeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics*, *39*(4), 457–466.

Weeks, A. R., Heinze, D., Perrin, L., Stoklosa, J., Hoffmann, A. A., van Rooyen, A., Kelly, T., & Mansergh, I. (2017). Genetic rescue increases fitness and aids rapid recovery of an endangered marsupial population. *Nature Communications*, *8*(1), 1071.

Weeks, A. R., Sgro, C. M., Young, A. G., Frankham, R., Mitchell, N. J., Miller, K. A., Byrne, M., Coates, D. J., Eldridge, M. D. B., Sunnucks, P., Breed, M. F., James, E. A., & Hoffmann, A. A. (2011). Assessing the benefits and risks of translocations in changing environments: a genetic perspective. *Evolutionary Applications*, *4*(6), 709–725.

Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2017). Direct determination of diploid genome sequences. *Genome Research*, *27*(5), 757–767.

Wendel, J. F. (2015). The wondrous cycles of polyploidy in plants. *American Journal of Botany*, *102*(11), 1753–1756.

Weston, L. M., Mattingly, K. Z., Day, C. T. C., & Hovick, S. M. (2021). Potential local adaptation in populations of invasive reed canary grass (Phalaris arundinacea) across an urbanization gradient. *Ecology and Evolution*, *11*(16), 11457–11476.

Wet, J. M. J., & Harlan, J. R. (1970). *Bothriochloa intermedia* — a taxonomic dilemma. *Taxon*, *19*(3), 339–340.

Whalley, J. A. T. R. (1977). Basal area, dry weight and size distribution of individual plants of some. *australian rangeland society*.

Whitlock, M. C. (2008). Evolutionary inference from QST. *Molecular Ecology*, *17*(8), 1885–1896.

Whitlock, M. C., & Guillaume, F. (2009). Testing for spatially divergent selection: comparing QST to FST. *Genetics*, *183*(3), 1055–1063.

Wicker, T., & Keller, B. (2007). Genome-wide comparative analysis of copia retrotransposons in *Triticeae*, rice, and *Arabidopsis* reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. *Genome Research*, *17*(7), 1072–1081.

Wiens, J. J. (2016). Climate-related local extinctions are already widespread among plant and animal species. *PLoS Biology*, *14*(12), e2001104.

Williams, A. V., Nevill, P. G., & Krauss, S. L. (2014). Next generation restoration genetics: applications and opportunities. *Trends in Plant Science*, *19*(8), 529–537.

Williams, C. F., Ruvinsky, J., Scott, P. E., & Hews, D. K. (2001). Pollination, breeding system, and genetic structure in two sympatric *Delphinium* (Ranunculaceae) species. *American Journal of Botany*, *88*(9), 1623–1633.

Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). *Limits to the Adaptive Potential of Small Populations*. https://doi.org/10.1146/annurev.ecolsys.37.091305.110145

Winter, D., Lee, K., & Cox, M. (2020). *Pafr: Read, Manipulate and Visualize "Pairwise mApping Format" Data. R package version 0.0.2*. https://CRAN.R-project.org/package=pafr (Original work published 2020)

Woodhouse, M. R., Schnable, J. C., Pedersen, B. S., Lyons, E., Lisch, D., Subramaniam, S., & Freeling, M. (2010). Following tetraploidy in maize, a short deletion mechanism removed genes preferentially from one of the two homeologs. *PLoS Biology*, *8*(6), e1000409.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(33), 13875–13879.

Wootton, J. T., & Pfister, C. A. (2013). Experimental separation of genetic and demographic factors on extinction risk in wild populations. *Ecology*, *94*(10), 2117–2123.

Wortemann, R., Herbette, S., Barigah, T. S., Fumanal, B., Alia, R., Ducousso, A., Gomory, D., Roeckel-Drevet, P., & Cochard, H. (2011). Genotypic variability and phenotypic plasticity of cavitation resistance in *Fagus sylvatica* L. across Europe. *Tree Physiology*, *31*(11), 1175–1182.

Wright, J. W., & Stanton, M. L. (2006). Local adaptation to serpentine and non-serpentine soils in Collinsia sparsiflora. *Evolutionary Ecology*. http://www.evolutionary-ecology.com/abstracts/v08/1761.html

Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics*, *15*(4), 323–354.

Wright, S. (1984). Evolution and the Genetics of Populations, *Volume 4: Variability Within and Among Natural Populations*. University of Chicago Press.

Wu, H., Wang, S., Wei, X., & Jiang, M. (2019). Sensitivity of seed germination to temperature of a relict tree species from different origins along latitudinal and altitudinal gradients: implications for response to climate change. *Trees*, *33*(5), 1435–1445.

Wulff, R. D., & Bazzaz, F. A. (1992). Effect of the parental nutrient regime on growth of the progeny in *Abutilon theophrasti* (malvaceae). *American Journal of Botany*, *79*(10), 1102–1107.

Wu, S., Han, B., & Jiao, Y. (2020). Genetic contribution of paleopolyploidy to adaptive evolution in angiosperms. *Molecular Plant*, *13*(1), 59–71.

Xiao, S. H. U., Yang, X., & Yang, Z. (2012). Variation in Seed and Seedling Traits among Fifteen Chinese Provenances of *Magnolia officinalis*. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, *40*(2), 274–283.

Yao, Y., Carretero-Paulet, L., & Van de Peer, Y. (2019). Using digital organisms to study the evolutionary consequences of whole genome duplication and polyploidy. *PloS One*, *14*(7), e0220257.

Yeaman, S., & Whitlock, M. C. (2011). The genetic architecture of adaptation under migration-selection balance. *Evolution; International Journal of Organic Evolution*, *65*(7), 1897–1911.

Yin, M., Wang, Y., Zhang, L., Li, J., Quan, W., Yang, L., Wang, Q., & Chan, Z. (2017). The Arabidopsis Cys2/His2 zinc finger transcription factor ZAT18 is a positive regulator of plant tolerance to drought stress. *Journal of Experimental Botany*, *68*(11), 2991–3005.

Young, A., Boyle, T., & Brown, T. (1996). The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, *11*(10), 413–418.

Yu, Prakash, & Whalley. (2000). Comparative reproductive biology of the vulnerable and common grasses in *Bothriochloa* and *Dichanthium*. *: Systematics and Evolution …*. https://books.google.com/books?hl=en&lr=&id=nocGdY9MViMC&oi=fnd&pg=PA307&dq=Comparative+reproductive+biology+of+the+vulnerable+and+common+grasses+in+Bothriochloa+and+Dichanthium+In+Grasses+Systematics+and+Evolution&ots=I-49IgAVO2&sig=e2gacAVU_uvKaFOjj5DLg3a-13Y

Zaharia, M., Bolosky, W. J., Curtis, K., Fox, A., Patterson, D., Shenker, S., Stoica, I., Karp, R. M., & Sittler, T. (2011). Faster and more accurate sequence alignment with SNAP. In *arXiv [cs.DS]*. arXiv. http://arxiv.org/abs/1111.5572

Zhai, X.-Y., Guo, Y.-X., Hou, F.-J., Liu, Y., Ma, D.-T., Wang, W., & Yan, X.-B. (2015). Population genetic structure and germplasm conservation of *Stipa purpurea* on Qinghai-Tibetan Plateau under grazing. *Biochemical Systematics and Ecology*, *62*, 51–57.

Zhang, J. (2003). Evolution by gene duplication: an update. *Trends in Ecology & Evolution*, *18*(6), 292–298.

Zhang, J., Zhang, X., Tang, H., Zhang, Q., Hua, X., Ma, X., Zhu, F., Jones, T., Zhu, X., Bowers, J., Wai, C. M., Zheng, C., Shi, Y., Chen, S., Xu, X., Yue, J., Nelson, D. R., Huang, L., Li, Z., … Ming, R. (2018). Allele-defined genome of the autopolyploid sugarcane *Saccharum spontaneum* L. *Nature Genetics*, *50*(11), 1565–1573.

Zhang, L., Wu, S., Chang, X., Wang, X., Zhao, Y., Xia, Y., Trigiano, R. N., Jiao, Y., & Chen, F. (2020). The ancient wave of polyploidization events in flowering plants and their facilitated adaptation to environmental stress. *Plant, Cell & Environment*, *43*(12), 2847–2856.

Zhang, T., Hu, Y., Jiang, W., Fang, L., Guan, X., Chen, J., Zhang, J., Saski, C. A., Scheffler, B. E., Stelly, D. M., Hulse-Kemp, A. M., Wan, Q., Liu, B., Liu, C., Wang, S., Pan, M., Wang, Y., Wang, D., Ye, W., … Chen, Z. J. (2015). Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. *Nature Biotechnology*, *33*(5), 531–537.

Zhang, X., Zhang, S., Zhao, Q., Ming, R., & Tang, H. (2019). Assembly of allele-aware, chromosomal-scale autopolyploid genomes based on Hi-C data. *Nature Plants*, *5*(8), 833–845.

Zhang, X.-Y., Hu, C.-G., & Yao, J.-L. (2010). Tetraploidization of diploid Dioscorea results in activation of the antioxidant defense system and increased heat tolerance. *Journal of Plant Physiology*, *167*(2), 88–94.

Zhao, K., Tung, C.-W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., Norton, G. J., Islam, M. R., Reynolds, A., Mezey, J., McClung, A. M., Bustamante, C. D., & McCouch, S. R. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. *Nature Communications*, *2*, 467.

Zhao, Y., Liu, Z., & Wu, J. (2020). Grassland ecosystem services: a systematic review of research advances and future directions. *Landscape Ecology*, *35*(4), 793–814.

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* , *28*(24), 3326–3328.