Germplasm exploration and phenotyping in *Trifolium* species for the improvement of agronomic traits and abiotic stress tolerance

---

A thesis
submitted in partial fulfilment
of the requirements for the Degree of
Doctor of Philosophy
at
Lincoln University
by
Lucy Marie Egan

Lincoln University
2020

Abstract of a thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy.

Germplasm exploration and phenotyping in *Trifolium* species for the improvement of agronomic traits and abiotic stress tolerance

by

Lucy Marie Egan

Trifolium is the most important pastoral legume genus for temperate agriculture. However, there has been little effort into characterising variation in *Trifolium* accessions. A series of studies utilising pedigree and genomic data were conducted to analyse variation in *Trifolium* accessions in the Margot Forde Germplasm Centre and in white clover breeding populations. Pedigree analysis experiments were designed to develop pedigree maps, calculate inbreeding and kinship coefficients, calculate the effective number of founders and identify influencing founders and ancestors for *Trifolium repens* (white clover), *Trifolium pratense* (red clover), *Trifolium arvense, Trifolium ambiguum, Trifolium dubium, Trifolium hybridum, Trifolium medium, Trifolium subterraneum* and *Trifolium repens* x *Trifolium occidentale* interspecific hybrids.

One of the earliest parental white clover accessions identified was 'North Canterbury Type I' from 1941. The Type 1 phenotype influenced the population structure of all white clover accessions in the genebank. The relatedness and inbreeding coefficients revealed distinct germplasm pools formed across time that are of interest to pre-breeding efforts. One germplasm pool contained the majority of commercial white clover cultivars, of which they had similar phenotypes.

Inbreeding in red clover remained stable in the last three decades, and a relationship between inbreeding and new introductions into the collection was found (r = 0.62). Founding accessions were classed into three phenotype groups; English Broad, English Giant Hybrid and Cotswald Broad. The first synthetic form was identified from parental accessions from English Broad and English Giant Hybrid.

Within the minor *Trifolium* species, kinship levels remained below 8%, and no inbreeding was found. *T. ambiguum* and *T. medium* had the highest cumulative kinship across the decades. The Australian cultivar 'Monaro' had a strong influence over all of the studied accessions in *T. ambiguum*.

Two genome-wide associations studies were designed to test for phenotype-genotype associations. The Mainstay panel had 242 white clover half-sib families, and the Genetic Gain panel had 80 white clover cultivars. The Mainstay panel had stolon density and growth score phenotype data. In contrast, the Genetic Gain panel had yield, clover content, seed yield, flowering duration, the peak number of flowers, leaf size, seed per head, leaf marking, cyanogenesis, normalised transpiration rate and the fraction of transpirable soil water phenotype data. Population structure was identified in the Mainstay panel (k=2), but no subpopulation divergence was present in the Genetic Gain panel. Both panels had rapid linkage disequilibrium decay. A marker-trait association was identified in the Genetic Gain panel for cyanogenesis. However, when the marker was tested against the marker dosage, it was not significant. The results of the genome-wide association studies emphasise the highly variable and heterozygous nature of white clover populations and cultivars.

Keywords: *Trifolium*, white clover, red clover, breeding, forage, pedigree, relatedness, inbreeding, founders, germplasm, pre-breeding, ancestors, transpiration, drought, white clover, normalised transpiration rate, fraction of transpirable soil water, genome-wide association, population structure.

Acknowledgements

I truly believe that it takes a village to get a PhD student across the line. This PhD research and thesis would not have been possible without many people, who I am extremely thankful to.

Firstly, I would like to thank my supervisors, Dr Rainer Hofmann (Lincoln University) and Dr Valerio Hoyos-Villegas (McGill University, formerly AgResearch) for their immense knowledge, patience, guidance and support throughout my project. Their doors were always open, and nothing was ever a problem. Even when Valerio moved to Canada, his support and speed at replying never waned.

I am very grateful for Dr Chris Winefield (Lincoln University) and Dr Ross Bicknell (Plant and Food Research) for teaching me the fundamentals of plant breeding, both practically and in theory. Without their passion and commitment to teaching, I do not believe I would be in this position.

Throughout this PhD, I had help from many incredible staff at AgResearch. Thank you to Dr Aurelie Laugraud, Dr Ken Dodds, Dr Rachael Ashby, Dr Abdul Baten, Dr Ruy Jauregui, Paul MacLean and Dr Chikako van Koten for their help with bioinformatics and statistics. Thank you to Angus Heslop and Tony Hilditch for their help, practical advice and discussions regarding field trials. Thank you to Dr Andrew Griffiths and Dr Marty Faville for their advice and knowledge with the genome-wide association studies. Thank you to Dr Marcelo Carena and Dr Tony Conner for their continuous support with my project. I would like to thank AgResearch for allowing me to complete my PhD at the organisation. A special thanks to Dr Jeanne Jacobs for being a fantastic mentor and support person, and for always empowering women in science.

I would like to thank Keith Widdup for awarding me an AgResearch summer studentship in 2014 and spiking my interest in plant breeding. His passion and encouragement have led me to pursue plant breeding as a career.

Thank you to the Lincoln University staff who have provided me with assistance; Dr Dean O'Connell for statistics advice and Stephen Stilwell for help in the lab.

Thank you to Dr Paul Shaw from the James Hutton Institute for help with his programme, Helium. Without him, there would be no pedigree maps.

Thank you to the following companies and trusts for the financial support and awards throughout the PhD: the Foundation of Arable Research postgraduate scholarship, the Manning Seed award, the John W & Carrie McLean Trust scholarship, the Beef and Lamb NZ Generic Ag scholarship, the Graduate Women Canterbury Trust travel award and the Royal Society of New Zealand Canterbury branch travel grant. I would like to acknowledge the financial support from Pastoral Genomics+ and AgResearch.

Finally, I wish to thank my close friends and family, parents, sister and brother-in-law, for their love and support. I dedicate this thesis to my nephew, Jack, in hopes that this will show him that anything is possible if you set your mind to it, and that growth is not possible in your comfort zone.

Table of Contents

List of Tables

List of Figures

Abbreviations

AHS	Among half-sibs
ANOVA	Analysis of variance
AWHS	Among and within half-sibs
BC_1	First backcross
BC_1F_2	First backcross second generation
BC_2	Second backcross
BLUP	Best linear unbiased predictor
COP	Completeness of parentage
FDR	False discovery rate
FTSW	Fraction of transpirable soil water
FTSWc	Fraction of transpirable soil water critical threshold
GBS	Genotyping-by-sequencing
GM	Genetic modification
GWAS	Genome-wide association study
GxE	Genotype x environment
HCN	Hydrogen cyanide
HWE	Hardy-Weinberg equilibrium
ISH	Interspecific hybrid
KASP	Kompetitive allele-specific polymerase chain reaction
LD	Linkage disequilibrium
MAF	Minor allele frequency
MAS	Marker-assisted selection
MFGC	Margot Forde Germplasm Centre
NGS	Next-generation sequencing
NPGS	National Plant Germplasm System
NTR	Normalised transpiration rate
PC	Principal component
PSC	Pseudo-self-compatibility
QQ	Quantile-Quantile
QTL	Quantitative trait loci
SNP	Single nucleotide polymorphism

Chapter 1
Introduction

1.1 The *Trifolium* genus

Trifolium is the most important pastoral legume genus to New Zealand agriculture (Caradus, et al., 1989a, Caradus, et al., 1996c). Approximately 250 species comprise *Trifolium*, but only ten *Trifolium* species are economically important in agricultural systems (Zohary, et al., 1984). White clover (*Trifolium repens*) is the most important pastoral legume in temperate regions of the world and is the most common pastoral legume in New Zealand (Caradus, et al., 1996c, Haynes, 1980). Economically, white clover is vital to New Zealand pastures (Caradus, et al., 1996c), and the economic benefit is reflected in the large investment into research and breeding. Breeding programmes for white clover have been present in New Zealand since the 1920s and have been successful in producing a broad portfolio of cultivars (Bouton, et al., 2005).

Red clover (*Trifolium pratense*) is less prominent than white clover in New Zealand and is often used as silage and hay or, is mixed with white clover in a pastoral system (Cassileth, 2010, Kemp, et al., 1999). The major benefit of red clover compared to white clover is the increased performance in areas prone to drought (Riday, 2010, Vaseva, et al., 2011). Red clover breeding has been present in New Zealand since the 1930s (Grasslands Division, 1971, Wratt, et al., 2015) and the introduction of increased ploidy has been instrumental in developing cultivars with increased performance (Taylor, et al., 1985, Taylor, et al., 1996b).

While white and red clover are the two most dominant species in temperate pastures, other *Trifolium* species such as *T. ambiguum, T. arvense, T. dubium, T. hybridum, T. medium* and *T. subterraneum* are often used for research purposes. They have been a source of variation for pre-breeding and interspecific hybridisation with white clover (Ellison, et al., 2006). Interspecific hybrids (ISH) between *T. repens* and *T. occidentale*, *T. uniflorum* and *T. ambiguum* have been successful in developing hybrids with superior performance to white clover in marginal climates (Abberton, et al., 1998, Nichols, et al., 2014c, Nichols, et al., 2015).

1.2 Germplasm centres and pre-breeding

Germplasm centres are vital to plant breeding efforts worldwide. They are defined as a centre designed to conserve large amounts of genetic diversity by storing germplasm (Williams, 2000). The germplasm contained, usually seeds, are from populations that are genetically diverse and are beneficial for plant breeding and conservation programmes (Ghimiray, et al., 2017). The Margot Forde Germplasm Centre (MFGC) is New Zealand's national genebank for grassland plants and contains seeds from 1800 species collected from over 100 countries (Egan, et al., 2019a, Egan, et al., 2019b,

Egan, et al., 2020). There are over 74,000 *Trifolium* accessions held globally in genebanks, and the MFGC contains approximately 9% of global *Trifolium* accessions (FAO, 2010).

Wild germplasm stored in genebanks are often used to introduce new traits into populations through hybridisation. The importance of wild and minor species to widen the gene pool has been reflected in increased efforts to collect and exchange germplasm around the world. However, to effectively utilise wild germplasm, characterisation of the germplasm is beneficial. Pre-breeding is defined as all activities designed to identify useful traits or genes from unadapted germplasm that can be integrated into breeding programmes to produce new cultivars (Sharma, 2017, Singh, et al., 2018). Without screening wild germplasm, it is of limited value in a breeding programme, and the lack of characterisation of germplasm has been attributed to the low utilisation rates of wild germplasm by breeders (Acquaah, 2012).

1.2.1 Plant variation in breeding programmes

The success of plant breeding activities is dependent on the variation present in the population. The imminent threat of changing climatic conditions and the need to produce cultivars to feed an ever-growing world has increased the need to ensure variation is present in breeding populations (Acquaah, 2012).

Population structure and quantitative metrics concerning variation present in germplasm collections in major crops have been widely reported (Brazauskas, et al., 2011, Briggs, et al., 2006, Casler, 1998, Gorjanc, et al., 2016, McNally, et al., 2009). However, the variation within populations and genebanks in *Trifolium* is largely unknown. Variation in *Trifolium* and other forage species is often generated through the highly heterozygous and outcrossing nature of the species. Yet, the slow rate of genetic gain in forages such as perennial ryegrass (*Lolium perenne*) (Harmer, et al., 2016), Italian ryegrass (*Lolium multiflorum* Lam) (Vanwijik, et al., 1991), alfalfa (*Medicago sativa* L.) (Hill Jr, et al., 1988), and more recently in white clover (Hoyos-Villegas, et al., 2019), indicates that increasing knowledge about variation, relatedness and population structure in *Trifolium* species could aid in increasing the rate of genetic gain and enhance environmental adaptation.

Pedigree analysis

Pedigree analysis can involve generating a pedigree map, deriving quantitative metrics and visualising population structure from the accessions (Holland, et al., 2003). Pedigree maps allow a visual look at the breeding on a population level, allowing subpopulations to be identified. Phenotypic and genotypic information can be added to pedigrees to add depth and knowledge to aid selections (Breseghello, et al., 2013). Population structure can visualise clustering patterns and can be exploited in future breeding decisions by breeding within and between clusters (Priolli, et al., 2010). Where available, molecular tools are utilised in mixed models alongside the pedigrees (Janick, 2003).

Genome-wide association study

A genome-wide association study (GWAS) is a study that observes a set of genetic variants and rapidly scans markers across the complete set of genomes to find genetic variants that are associated with a trait of interest. In recent years, GWAS has become applicable due to the establishment of genotyping pipelines and large-scale genome-wide single nucleotide polymorphism (SNP) resources (Byrne, et al., 2013, Hand, et al., 2012). As well as identifying marker-trait associations, GWAS can give a genome-level view of the variation present in the study population. In model plants, GWAS has provided knowledge about the allelic architecture of simple and complex traits and has accounted for observed phenotypic variation in some species (Atwell, et al., 2010, Biazzi, et al., 2017, Manolio, et al., 2009, Sakiroglu, et al., 2017, Sakiroglu, et al., 2012, Vineis, et al., 2010).

1.3 Gaps in knowledge

In *Trifolium*, there has been minimal focus on the levels of variation present in breeding populations (Jahufer, et al., 2013a) and population-specific patterns that have occurred due to artificial selection. The only published studies on pedigree analysis in *Trifolium* species are from this thesis (Egan, et al., 2019a, Egan, et al., 2019b, Egan, et al., 2020), and there is only one published GWAS study in white clover (Inostroza, et al., 2018). Pedigree analysis allows insight into the effect that breeding decisions and selection have had on population structure and relatedness. At the same time, GWAS will show the variation on a genomic level and could identify marker-trait associations.

1.4 Objectives

The principal objective of this thesis was to analyse variation in *Trifolium* species. The following six objectives were identified to achieve this objective:

Variation within a genebank:

1. Identify population-specific patterns and calculate relatedness and inbreeding coefficients for white clover, red clover and related *Trifolium* species used in breeding programmes.

2. Identify implications of founding accessions on *Trifolium* breeding programmes.

3. Investigate variation present in *Trifolium* related species collections that can be utilised in pre-breeding decisions.

Phenotypic variation between cultivars:

4. Establish the fraction of transpirable soil water threshold for white clover cultivars.

Genomic tools to assess variation and phenotype-genotype relationships:

5. Investigate the population structure of white clover populations through genomic analyses.

6. Use a genome-wide association study to identify marker-trait associations in white clover populations.

1.5 Outline of the thesis

The structure of the thesis is shown in Figure 1.1. Six experimental chapters, chapters 3, 4, 5, 6, 7 and 8, are presented to address the objectives of the thesis. The key findings are summarised in chapter 9 and the overall implications of the research are discussed.

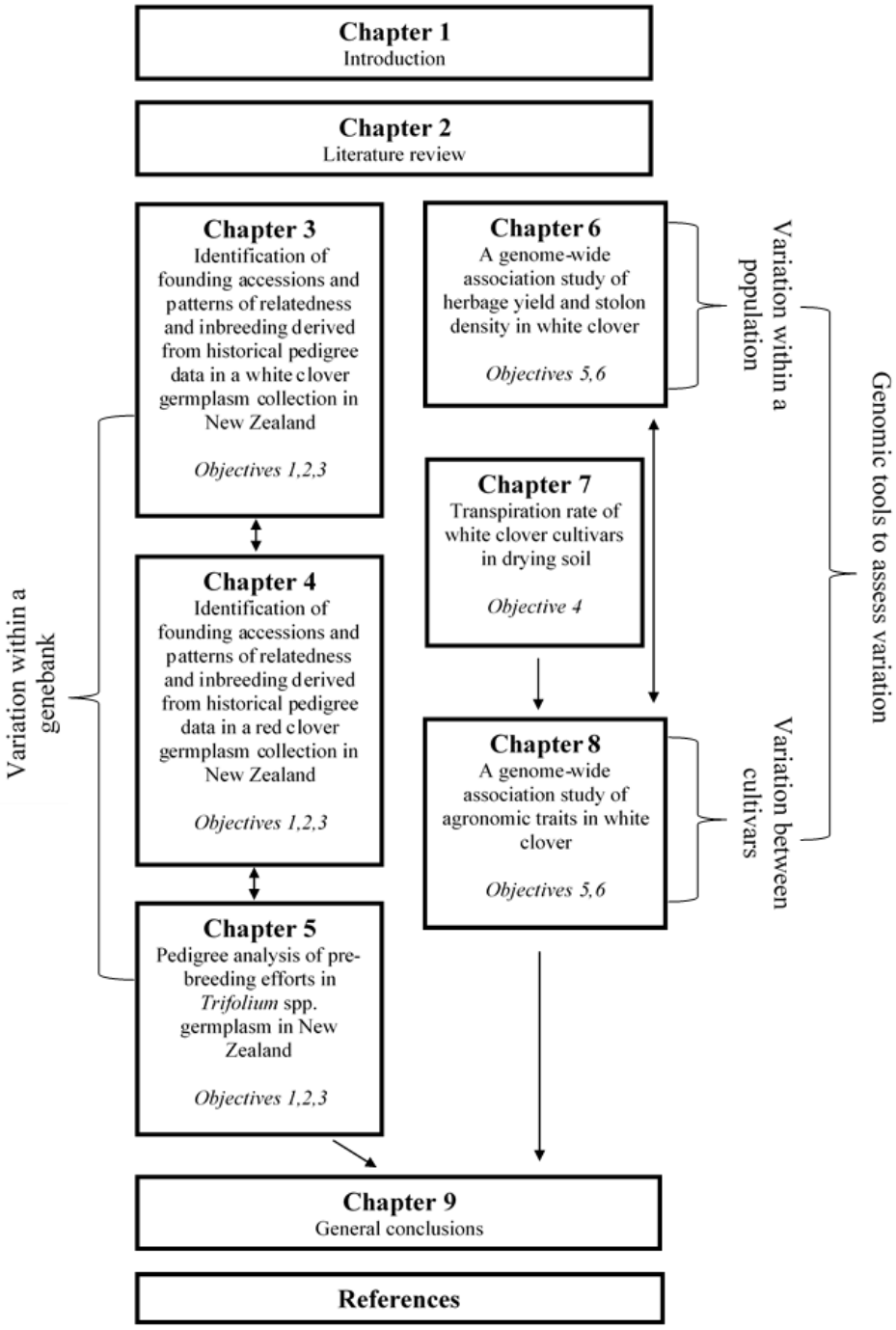

Figure 1.1 Flow diagram of the thesis structure showing the relationships between chapters.

Chapter 2

Literature review

This chapter has been prepared for submission to Crop and Pasture Science.

2.1 Introduction

Trifolium species are the most important and valuable forage legumes in the temperate world (Russel, et al., 1976). The genus *Trifolium* includes more than 250 species, many of which have agricultural importance (Zohary, et al., 1984). The most widely used species within the genus are white and red clover. White clover is the most critical pastoral legume in temperate zones of the world. It is grown widely throughout pastoral systems in Europe, western Asia, North America, Australia and New Zealand. It is of value to agricultural pastures due to its high nutritional value and quality, persistence, wide climatic range of growth, high seed production and ability to fix atmospheric nitrogen (Baker, et al., 1987). While white clover can be grown in a wide range of climates, it is mostly grown in temperate mild to cold temperate climates (Frame, et al., 1986).

White clover is an allotetraploid (2*n*=4*x*=32) perennial legume and exhibits disomic inheritance. The genome size is compact (1C = 1093Mb) (Bennett, et al., 2011). White clover is a recent allopolyploid, estimated to have arisen 13,000 to 130,000 years ago through the hybridisation of two diploid ancestors; *Trifolium occidentale* and *Trifolium pallescens* (Ellison, et al., 2006, Griffiths, et al., 2019, Williams, et al., 2012). The mating system is highly outcrossing with a gametophytic self-incompatibility system, which develops highly heterozygous populations (Abberton, 2007, Williams, et al., 2012).

Red clover is native to Europe, western Asia and north-western Africa. It is grown widely as a fodder crop that is used as silage and hay. Red clover is a diploid (2*n*=2*x*=14) perennial legume with a genome size of ~420Mb (De Vega, et al., 2015, Ištvánek, et al., 2014). Like white clover, red clover is almost completely self-sterile but when outcrossed, produces highly variable populations. Red clover cultivars can be diploid or tetraploid, and the tetraploid cultivars have been known to outperform the diploid cultivars in some agronomic traits (Taylor, 2008).

Trifolium subterraneum, known as subterranean clover, is an annual forage legume species of clover native to north-western Europe and western Asia. Of all the annual clovers, subterranean clover has the highest contribution to livestock feed production (Kaur, et al., 2017). Subterranean clover is a mostly self-pollinating, diploid (2n = 16) species (Ghamkhar, et al., 2012, Hirakawa, et al., 2016, Kaur, et al., 2017).

Trifolium ambiguum, commonly known as 'Caucasian' clover, is a species native to Asia. *T. ambiguum* can be diploid (2n = 16), tetraploid (2n = 32) or hexaploid (2n = 48) (Bryant, 1974, Taylor, et al., 1997). Although the ploidy of the species affects some traits such as flowering date and persistence, the yield is unaffected (Bryant, 1974, Dear, et al., 1985).

Trifolium arvense, Trifolium dubium, Trifolium hybridum and *Trifolium medium* are four *Trifolium* species which have a minor presence in pastoral systems and research. *T. arvense*, commonly known as rabbitfoot clover, is an annual clover that is native to Europe and Western Asia (Pritchard, et al., 1988). *Trifolium dubium*, known as 'Suckling clover', is an allotetraploid (2n = 4x = 30) clover that is native to Europe (Bulińska-Radomska, 2000). It arose from a cross of *T. campestre* and *T. micranthum* (Hedlund, et al., 2003). *T. hybridum*, commonly known as 'Alsike clover', is a clover that originates from Europe and has established throughout temperate regions of the world (Williams, 1951). *Trifolium hybridum* is highly self-sterile (Williams, 1951). *T. medium*, commonly known as 'zigzag clover', is a native European perennial species with the ploidy of 2n=10x=80 (Isobe, et al., 2002, Merker, 1984). It is similar in appearance to red clover but with narrower leaflets and no white leaf markings. *Trifolium medium* is known to have long persistence (Choo, 1988).

2.2 Role of *Trifolium* species in New Zealand pastoral systems

2.2.1 White clover

White clover is often described as the base legume for New Zealand's pastoral sector. Annually, 1000-1200 tonnes of white clover seed are sold in New Zealand, and 4500 tonnes are exported around the world. New Zealand has the highest global export share of white clover seed (57.5%) (Rattray, 2005). White clover is commonly used in a mixture with grass and is grazed *in situ*. It is tolerant to a range of grazing systems, including dairy, sheep and beef, cattle and deer, and is favourable in grazing systems due to the high feed value (Frame, et al., 1986). Traditionally, white clover was not used in hay or silage. The brittleness of the leaves, lack of bulk production and problems with producing well-fermented silage has since been overcome by wilting, chopping and the use of acid additives (Baker, et al., 1987).

The nitrogen-fixing capability of white clover is favourable in a sward as it reduces the need for synthetic fertilisers for the companion grasses. Crush (1987) estimated that white clover has the potential to fix 600-700kg N/ha/year. However, varying abiotic conditions can lower nitrogen fixation rates. White clover fixes approximately 1.57 million tonnes of nitrogen annually, contributing approximately $1.49 billion to the New Zealand economy (Caradus, et al., 1996c).

Brown, et al. (2003) summarised the main constraints of white clover production, being (i) competition with grasses, (ii) competition from pests, (iii) competition from nitrogen fertiliser, (iv) moisture availability, (v) persistence in drought, and (vi) lower temperature thresholds. Warming global temperatures signifies that performance under long-term drought is one of the most severe

constraints to white clover (Brown, et al., 2003, Macfarlane, et al., 1990b, Sheath, et al., 1990). The ideal growth temperature for white clover is 20-24˚C, and when the temperatures are not optimal, production decreases (Harris, et al., 1985). The poor performance under drought has significant effects on production. Studies show that white clover cultivars have fluctuating herbage yield every year in response to environmental stresses (Jahufer, et al., 2002, Jahufer, et al., 2013a) and has been attributed to poor survival through the summer drought conditions (Gillard, et al., 1989, Robinson, et al., 1976).

2.2.2 Red clover

Red clover is a perennial clover with a large taproot and high feed value. Red clover is not as prevalent as white clover in New Zealand pastoral systems, although approximately 100-150 tonnes of red clover seed is sold in New Zealand annually. Worldwide, it is primarily used as a fodder crop as silage and hay. But when used in a pastoral system, red clover is often mixed with white clover in pasture mixes (Cassileth, 2010, Kemp, et al., 1999). Red clover does not persist or perform well in intensive grazing systems and is better suited for dry summer areas where less intensive grazing systems are present (Kemp, et al., 1999).

Red clover offers unique advantages when compared to white clover: (i) faster establishment time and better performance in dry summer environments (Brown, et al., 2005, Moot, et al., 2000), (ii) increased water extraction in water-limited areas owing to the deep taproot (Taylor, et al., 1996a), (iii) increased tolerance to pasture pests (Gerard, et al., 2017), and (iv) improved nitrogen partitioning when consumed by livestock (Sullivan, et al., 2006, Van Ranst, et al., 2011). However, the biggest inhibitor of red clover performance in pastoral systems is the lack of persistence (Ford, et al., 2011). The longevity of red clover is generally 2-3 years, and when sub-optimal pasture conditions are present, persistence continually decreases (Hyslop, et al., 1999). Red clover persistence is reduced with frequent, hard grazing. Continuous grazing reduces the carbohydrates available in the taproot and increases the vulnerability of the plant to disease. The survival and performance of the taproot is a vital aspect of the survival and performance of the plant as a whole (Brock, et al., 2003, Smith, et al., 1985).

Pests and diseases are most damaging during red clover establishment. As the sward increases in cover, red clover is more competitive and less susceptible to infection from exposure in the sward (Frame, 2019). The predominant crown pathogen (*Sclerotinia trifoliorum*) and root pathogens (*Fusarium* spp. and *Rhizoctonia* spp.) are the major diseases to red clover. At the same time, weevils and nematodes are major pests (Hyslop, et al., 1999). Often, wounding of the plants from grazing animals leaves the plant open and susceptible to infection and damage. However, usually, the late-flowering cultivars have higher persistence and lead to less damage from pests and diseases early in the establishment. There is a positive correlation between ploidy levels and increased resistance to pests, showing variation amongst red clover cultivars (Hay, et al., 1989).

A common animal health issue from grazing red clover is bloating. Bloating is often more common in cattle than sheep. High levels of protein gas form in the stomach of the animal and when the gas levels become more elevated than the animals' ability to expulse the gas, bloating occurs (Majack, et al., 2003). Pasture mixes containing grass and grazing management plans are utilised to reduce the risk of bloating (Frame, et al., 1986).

2.2.3 Other *Trifolium* species

Many *Trifolium* species are underutilised, or their use is yet undetermined in agricultural systems (Maxted, et al., 2001). The importance of minor species within the *Trifolium* genus is now recognised in (i) growth in adverse conditions, (ii) research to improve the major species, or (iii) hybridisation with major *Trifolium* species.

Trifolium subterraneum and *T. ambiguum* are two species that have become popular to use in drought-prone pastoral systems of Australasia and North America. *Trifolium subterraneum*, also known as sub clover, is a species that is highly utilised in Australia and to a lesser extent in New Zealand. *Trifolium subterraneum* performs well in drought conditions as it buries the seeds in the soil, allowing seed development to happen underground (Suckling, et al., 1983, Widdup, et al., 2000). *Trifolium ambiguum*, also known as Caucasian or Kura clover, has a large root system and is favourable in agricultural systems that are exposed to climate extremes (Bryant, 1974, Maxted, et al., 2001). *Trifolium ambiguum* has also been used to hybridise with white clover (Ellison, et al., 2006, Williams, 2014). The *T. repens* x *T. ambiguum* ISHs have been successful in producing progeny that have advantageous root characteristics (Abberton, et al., 1998), and similar forage quality (Marshall, et al., 2004) and nitrogen fixation ability to white clover (Abberton, et al., 2000).

Trifolium arvense, T. dubium, T. hybridum and *T. medium* are four *Trifolium* species that have been used primarily for research to improve other *Trifolium* species. *Trifolium arvense*, is an annual winter clover that grows well in sandy or non-irrigated land (Pritchard, et al., 1988). The optimal environment in pastoral systems for *T. arvense* is short-lived and low fertility pastures; often in dry hill country environments (Taylor, 1985, White, et al., 1999). *Trifolium arvense* can survive for long periods under intense grazing but is often outcompeted under lighter grazing (Palmer, 1972). *Trifolium dubium*, is often found in the low fertility hill country (Caradus, et al., 1989b). *Trifolium hybridum*, is a short-lived perennial clover. It is adaptable to a wide range of conditions and has rapid establishment (Widdup, et al., 1994). It is often used in the hill country in the South Island of New Zealand for pasture and hay or silage (Taylor, 1985, Williams, 1951). *Trifolium medium*, is a long-lived perennial clover that prefers damp, acidic soils (Taylor, 1985). *Trifolium medium* has a similar phenotype to red clover (Taylor, 1985) and was introduced into North America as a contaminant in red clover seed. Although not common in commercial pastoral systems, *T. medium* has the potential for pasture and hay production (Taylor, et al., 1984).

The objective of this review chapter is to describe the role of genebanks in *Trifolium* breeding and what the characterisation of germplasm held in the genebanks could mean for plant breeding and pre-breeding efforts. The paper will include past breeding efforts, the importance of genebank and core collections, as well as possible avenues to characterise and utilise variation. Emphasis will be given to review research pertinent to New Zealand agriculture, but examples of research from other countries are provided as well. This paper aims to build on previous *Trifolium* reviews (Caradus, et al., 1989a, Caradus, et al., 1996c, Jahufer, et al., 2013b, Williams, 2014).

2.2.4 White clover breeding history in New Zealand

White clover improvement was initiated in New Zealand in the 1870s. Caradus, et al. (1989a) reviewed the advances of white clover breeding in New Zealand. The 1920s saw a continual supply of phosphate to New Zealand, allowing increased pasture production and stocking rate, increasing production and efficiency (Caradus, et al., 1989a). The 1930s focussed on enhancing productivity and persistence in existing cultivars (Woodfield, et al., 1994). The breeding objectives changed over time with environmental pressures. The 1950s focussed on primary physiological and morphological responses to environmental changes. Stolon density and growth of the stolons concerning the production and persistence of white clover became a focal point in the 1970s. In comparison, the 1980s had an emphasis on whole plant studies and the regeneration rates of white clover. This was combined with the 1990's, where different farm and pasture management practices, such as utilising cultivars with a specific leaf size, were incorporated into programmes (Caradus, et al., 1989a).

The leaf size of white clover; large-, medium- or small-leaved, generally determines the type of production system the cultivar would be utilised within. Large-leaved clover grows tall and upright and has thick stolons and robust roots. They are used frequently in dairy systems as they perform well in rotationally grazed pastures. Although very productive, they have fewer stolons, reducing the persistence in mixed swards. Medium-leaved clovers are the most robust type of clover, performing very well under a range of grazing management, except under very close continuous grazing (Baker, et al., 1987). Often, large- and medium-leaved cultivars are used together in pasture mixes and grazed on dairy pastures. Small-leaved clovers are low-growing with high numbers of leaves and thin, multi-branched stolons. The compact and low-growing morphology makes it difficult for the grazing animals to uproot the plants. Therefore, small-leaved clover has a high tolerance for rigorous defoliation and are often used in sheep grazing systems (Rattray, 2005).

The four main types of white clover found in New Zealand pastures were initially identified due to differences in variation. Type 1 was, 'New Zealand Wild White No.1'; a very productive, cyanogenic, medium-leaved perennial found in fertile old pastures. Type 2 was, 'New Zealand Wild White No. 2'; a perennial, with denser and smaller leaves and therefore less productivity than Type 1. Type 3 was, 'Ordinary New Zealand White'; a non-persistent clover with medium-sized leaves and moderate growth in the first year but poor growth after that. Type 4 was, 'Lax early-flowering New Zealand and

ordinary European'; a non-persistent, near-annual type, with small leaves and low productivity (Caradus, et al., 1989a). Type 1 was superior and led the way to certification and commercial production in 1930 (Caradus, et al., 1996c). A breeding programme was created to breed the commercialised Type 1 with pedigree New Zealand Certified Mother Seed. The populations were improved continuously until 1957 when the final selection was completed. In 1964 the lines were released as 'Grasslands™ Huia', and the cultivar is still utilised today (Caradus, et al., 1989a).

Since the release of 'Grasslands™ Huia', many white clover cultivars have been bred and evaluated. However, the breeding system of white clover presents some unique challenges to breeders. Allotetraploidy in white clover arose out of the hybridisation of two diploid ancestors; *T. pallescens* and *T. occidentale*. The polyploid and outcrossing nature of white clover has advantages and disadvantages (Comai, 2005). As white clover is an outcrossing species, the populations are highly variable both within and between populations, allowing performance in a wide range of environments (Hoglund, et al., 1985). However, a high recombination rate can also mean that it is a slow process in breeding towards specific traits of interest.

Breeding for multiple traits is a challenge in white clover. There is a negative correlation for the selection of leaf size and stolon density. Leaf size is used as a measurement of yield under optimal conditions, and stolon density is used as a measure of persistence. The negative relationship of leaf size and stolon density has attempted to be overcome by improving persistence and herbage yield through crossing small-leaved New Zealand ecotypes and large-leaved overseas germplasm into new populations for selection. This strategy has been successful and has developed cultivars such as 'Grasslands™ Demand' and 'Grasslands™ Sustain' (Caradus, et al., 1997a, Widdup, et al., 1989). A study by van den Bosch, et al. (1993) showed that selection for elite characteristics of root morphology resulted in a decrease of yield and persistence. Breeders have observed that after the first year, overseas material display poor productivity and persistence than New Zealand-adapted cultivars (Caradus, et al., 1989a). Together with improved farming practices, white clover cultivars with improved persistence have been developed (Caradus, et al., 1996a, Charlton, et al., 1989, Lee, et al., 1993, Rhodes, et al., 1979).

2.2.5 Red clover breeding history in New Zealand

Red clover breeding has been more prominent overseas than in New Zealand and is reflected in the abundance of the plant distribution globally. Red clover improvements have been focussed on yield and persistence (Smith, et al., 1985, Taylor, et al., 1996a). Many of the cultivars produced are early-flowering type and lack persistence, as they are most commonly used as a forage supplement crop rather than in a long-term grazing system (Taylor, et al., 1996a). Despite the lack of persistence, red clover is considered an important component in pastoral systems throughout New Zealand. The cultivars released from New Zealand are agronomically similar to overseas cultivars but are adapted to New Zealand conditions. Much like white clover, all of the red clover cultivars in New Zealand are

synthetics and have been created through the open pollination of multiple elite parents (Wratt, et al., 2015). The introduction of polyploidy has been successful in red clover and led to cultivars such as 'Grasslands™ Pawera' and 'Grasslands™ G27' (Rumball, et al., 1997, Wratt, et al., 2015).

The first red clover plants imported into New Zealand were from commercial seed companies in England. There has been a lack of recording of New Zealand ecotypes that had adapted from the imported material. The most popular cultivar of choice for New Zealand farmers was the 'Broad Red' type. However, in the second year of the pastoral system, the red clover population was weak and sparse. The Montgomery type was slower to establish than Broad Red but provided a much more persistent population. The 1920s in New Zealand saw the beginning of the trials to determine the best type of red clover to use for New Zealand pastures, and the differences between Broad Red and Montgomery were apparent. Montgomery proved to be more successful for New Zealand pastures and developed into a breeding programme in the 1930s. In 1937, the seed was certified and was classed as New Zealand Montgomery red clover. It was renamed as 'Grasslands™ Turoa' in 1964 (Wratt, et al., 2015).

The 1930s showed a good establishment of Broad Red clover and several breeding programmes were developed to increase density and yield. In 1946, the successful parents from these trials were certified to be named as the cultivar New Zealand Broad Red clover (Wratt, et al., 2015), only to be renamed to 'Grasslands™ Hamua' in 1964 (Levy, 1970). 'Grasslands™ Relish' is the most recent New Zealand red clover cultivar to market. Relish can persist in pastoral systems for three to four years; a significant increase than other red clover cultivars (Ford, et al., 2011).

The development of tetraploid red clover began overseas but was rapidly recognised as needed in New Zealand. In New Zealand, the development of tetraploid red clover began in 1954, and by 1972 'Grasslands™ Pawera' was commercialised. Phenotypically, this cultivar is described as a mix between 'Grasslands™ Turoa' and 'Grasslands™ Hamua' (Wratt, et al., 2015). Tetraploids ($2n$=28) have distinct advantages when compared to diploids (Taylor, et al., 1996b). The doubling of chromosomes interrupts the alleles that control self-fertility, and often tetraploids exhibit increased numbers of self-fertile plants compared to diploids (Drach, et al., 1986). Tetraploid red clover cultivars often outperform diploid cultivars in dry matter yield and disease resistance (Arseniuk, 1989, Joensson, 1985, Liatukas, et al., 2012, Taylor, et al., 1996a, Yamada, et al., 1990). Tetraploids are usually larger in most plant structures, including flowers and seeds (Nikovitz, 1985). However, tetraploids have lower seed production compared to diploids, so higher sowing rates are needed. The lowered seed yield of tetraploids has been a major limiting factor in the production of tetraploid cultivars (Taylor, et al., 1996b).

Breeding for persistence in red clover has proved challenging. The need for increased persistence was recognised in the 1930s and similar to white clover, breeding efforts initially focussed primarily on plant morphology. Cultivar development has progressed with breeding for increased 'creeping'

phenotypes and stoloniferous features to increase persistence (Williams, et al., 2007). Although there have yet to be substantial gains made from increased grazing persistence, there have been successful developments in increased persistence in silage systems (Ford, et al., 2011).

The most prevalent chemical finding in red clover has been the discovery of phytoestrogens (Taylor, et al., 1996a). Phytoestrogens are a plant-derived estrogen that is structurally and functionally similar to mammalian estrogens (Patisaul, et al., 2010). Many phytoestrogens are phenolic compounds that can be linked back to the isoflavones group, which is present in many legume species. Further research showed a strong link between high levels of phytoestrogens in red clover, to decreased efficiency in ewe fertility, specifically formononetin (Rumball, et al., 1997). Formononetin is the causative hormone that is related to reduced conception and ovulation rates in ewes. Ewes can display two types of infertility when exposed to formononetin; temporary or permanent. Grazing management has allowed better control over the exposure to formononetin (Shackell, et al., 1993a, Shackell, et al., 1993b). However, the more recent red clover cultivars have been bred with lower levels of phytoestrogen. Five breeding cycles are usually needed to reduce the formononetin to a safe level for the sheep (Wratt, et al., 2015).

2.2.6 Minor *Trifolium* species

Trifolium subterraneum and *T. ambiguum* are two clovers that are widely cultivated in other countries, showing the potential to introduce these species into systems in New Zealand (Abberton, 2007). However, the utilisation of minor *Trifolium* species in New Zealand is still low.

The majority of *T. subterraneum* breeding has been in Australia by improving local germplasm (Taylor, 1985). Francis, et al. (1970) summarised that the key breeding targets for *T. subterraneum* were low oestrogen, marker characteristics, maturity, burr burial, physiological seed dormancy, seed hardness, and pathogen and insect resistance. Similar to red clover, *T. subterraneum* has high levels of formononetin which is generally selected out early in the breeding programme (Nicholas, et al., 1981). Trials across New Zealand evaluated *T. subterraneum* lines that have improved persistence and dry matter production (Dodd, et al., 1995a, Dodd, et al., 1995b, Dodd, et al., 1995c). Recent studies have identified Australian cultivars that are best adapted to New Zealand dryland pastoral systems (Lucas, et al., 2015, Olykan, et al., 2018). However, more improvement is needed for New Zealand environments.

Resistance to pathogens have a strong breeding objective in *T. subterraneum* breeding programmes. Pathogens such as *Kabatiella caulivora* (clover scorch), *Uromyces trifolii-repentis* (rust), *Oidium* sp. (powdery mildew), and *Fusarium avenaceum* and *Pythium irregulare* (root rot) are common in subterranean clover. There has been some success in developing resistant lines (Nichols, et al., 2014a, You, et al., 2005a, You, et al., 2005b).

Trifolium subterraneum was the first annual *Trifolium* species to have a draft genome sequenced. As it is an annual, diploid (2n=16) species with a small genome (540Mbp), it has been an attractive species to use as a model (Hirakawa, et al., 2016). Using *T. subterraneum* as a model species to understand the genetics of traits of interest will provide a pathway to understanding traits in the more genetically complex species.

Trifolium ambiguum has the potential to become a major forage legume in New Zealand. *T. ambiguum* exists in diploid, tetraploid and hexaploid forms and is highly self-incompatible at all ploidy levels. The presence of valuable traits in *T. ambiguum* have made it an attractive resource for interspecific hybridization. Although incompatibility exists between ploidy levels, there have been interploidal hybrids produced (Kannenberg, et al., 1962). The significant strengths of *T. ambiguum* are the longevity and persistence under intense grazing pastoral systems. Breeding efforts thus far have focussed on major agronomic traits such as seed and forage yield, and flowering time, as well as the more complex traits of drought resistance and performance with *Rhizobium* strains (Taylor, et al., 1997).

Trifolium ambiguum has been a considerable source of variation to introgress into white clover. One of the major goals with the *T. repens* x *T. ambiguum* hybrids is to introgress the large root system of *T. ambiguum*, while keeping the agronomic performance of white clover. The biggest challenges in these hybrids are maintaining seed production, slow establishment and producing viable hybrids (Meredith, et al., 1995, Taylor, et al., 1997). *Trifolium ambiguum* could also be a potential source of virus resistance to white clover. Barnett, et al. (1975) reported that *T. ambiguum* showed resistance to a range of viruses including alfalfa mosaic, yellow bean mosaic, peanut stunt and white clover mosaic viruses. Although *T. ambiguum* shows promise to become a productive forage legume in New Zealand pastoral systems, white and red clover remain very popular in the farming community due to their continuing high performance (Taylor, 2008).

The reproductive biology of *T. arvense* allows it to self- and cross-pollinate (Palmer, 1972). Breeding and research into *T. arvense* has been limited. However, a study by Hancock, et al. (2012) used genetic modification (GM) to integrate the transcription factor, TaMYB14, from *T. arvense* into *T. repens*. TaMYB14 is involved in the regulation of proanthocyanidin (PA) biosynthesis in legumes. PAs are polyphenolic secondary metabolites in plants and are associated with providing defence against pathogens and herbivores (Aziz, et al., 2005, de Colmenares, et al., 1998, Dixon, et al., 2005). The GM clover can decrease methane emissions and reduce bloating in livestock (Hancock, et al., 2012).

Trifolium hybridum is a self-incompatible, highly outcrossing species. Cultivars are either diploid (2*n*=16) or tetraploid (2*n*=36). There have been limited breeding programmes with *T. hybridum*, but it has been shown that there is wide variability in agronomic traits, except for persistence. Furthermore, inbreeding in *T. hybridum* reduces persistence (Matthews, et al., 1951, Townsend, 1964). Townsend, et al. (1968) assessed both self- and cross-pollinated populations to measure the effect that selection

for persistence had on the outcrossed progeny. The outcrossed populations had more persistence, but the gain was not enough to continue with selections.

Trifolium medium is highly self-incompatible and has 2*n* chromosome numbers ranging from 64 to 80 (Quesenberry, et al., 1977). *Trifolium medium* has been involved in several breeding programmes, including hybrid programmes (Taylor, et al., 1984). A draft genome of *T. medium* has been assembled to accelerate breeding advancements in clover breeding (Dluhošová, et al., 2018). There have been attempts to produce *T. medium* × *T. repens* and *T. medium* × *T. pratense* hybrids, but they have been unsuccessful (Anderson, et al., 1974, Kazimierska, 1978). *Trifolium medium* has been successfully crossed with *T. sarosiense* to bridge the genetic gap between *T. medium* and *T. pratense* (Quesenberry, et al., 1977). The primary breeding target of the *T. pratense* × *T. medium* hybrid is to incorporate increased perenniality.

The main breeding programme for *T. hybridum* in New Zealand was for the development of 'G41' zigzag clover. This programme focussed more on seed-setting than agronomic vigour. However, G41 has 84 chromosomes, so may be predisposed to meiotic mutations (Rumball, et al., 2005). For *T. medium* to be used as a potential forage legume in New Zealand pastoral systems, more research into seed traits and the agronomic and management practices in high-country systems is needed (Daly, et al., 1987, Taylor, 2008).

2.3 *Trifolium* breeding methods and progress

Breeding of *Trifolium* species in New Zealand commenced in the early 1900s. Early scientists recognised the importance of various *Trifolium* species to farm productivity which aided in the rapid expansion of agricultural production to supply animal (mostly sheep) products to England and other parts of the world (Caradus, et al., 1989a). Large investment into the breeding of *Trifolium*, and the resulting volume of research, has placed the *Trifolium* genus as the most important pastoral legumes to New Zealand pastures.

The breeding methods used in the past for *Trifolium* have produced cultivars that perform in a broad range of climates and farming systems (Caradus, et al., 1996b). The breeding techniques prevailing the 1960s were based on increasing the performance of ecotypes and phenotypic selection. Recurrent phenotypic selection, introduced in the mid-1960s (Williams, 1987), is a method of population improvement through the cyclical selection of the best performing plants within and among families, generation after generation, after the population has the selected desired traits (Hallauer, 1992).

Hoyos-Villegas, et al. (2018) compared among half-sib (AHS) family selection and among and within half-sib (AWHS) family selection strategies using computer simulation. AWHS family selection was a superior strategy, especially in early selection cycles. In the first selection cycle, AWHS had 2.5% genetic gain compared to 1% for AHS. Hoyos-Villegas, et al. (2019) suggested that recurrent selection increased the rate of genetic gain of yield and persistence in cultivars.

Phenotypic selection methods have been the most common for population improvement and cultivar development in *Trifolium* breeding. However, the literature is conflicted about how successful these methods have been in increasing the rate of genetic gain. While Woodfield, et al. (1994) stated that the rate of genetic gain for white clover yield and percentage clover in the sward was 6% per decade, Woodfield (1999) suggested that the rate of genetic gain rate 1.49% per year. More recently, Hoyos-Villegas, et al. (2019) showed that the rate of genetic gain in white clover did not reach above 2% per decade as established by the national strategy to lift pastoral sector productivity. Similar to white clover, the rate of genetic gain in red clover spans a moderate range. Estimates of the annual rate of genetic gain for forage yield in red clover are from 0.21% to 1.39% (Riday, 2010, Tucak, et al., 2013).

The development of molecular techniques in forage breeding has increased over the last decade. A popular type of molecular marker in forages are SNPs. SNPs are the most common type of genetic variation and are an efficient way of detecting genetic variation. Because of their abundance, SNP markers can increase the resolution and power in the detection of quantitative trait loci (QTL) that control traits and if an association is found and can be converted into kompetitive allele-specific polymerase chain reaction (KASP) markers for marker-assisted selection (MAS) breeding programmes. The discovery of QTL for seed and vegetative properties and the development of linkage maps are the first step in the development of MAS in *Trifolium* breeding programmes (Barrett, et al., 2005, Barrett, et al., 2009, Faville, et al., 2012, Isobe, et al., 2009, Williams, et al., 2007). The development of cost-efficient reduced representation genotyping techniques such as genotyping-by-sequencing (GBS) has increased the use and adoption of markers. SNP markers simplify some of the difficulties that occur in deploying markers in outcrossing species (Brummer, 2013, Elshire, et al., 2011).

2.4 Importance of genebanks and germplasm exploration in forages

Wild relatives of common species are becoming increasingly important to conserve and store in genebanks (Williams, 2010). Hybridising species with wild relatives broadens the gene pool and increases the available variation; both of which have been beneficial for modern agricultural practices. The hybridisation of crop species with wild relatives can increase the rate of genetic gain. In forages, examples include the crossing of forage rape (*Brassica napus* L.) and turnip (*Brassica campestris* L. ssp. *Rapifera*) as an alternative to forage rape (Mackay, 1973), the release of KX2Hawaii, a *Leucaena* hybrid cultivar, for increased pest tolerance (Brewbaker, 2008) and more recently, the successful crossing of *Paspalum plicatulum* and *Paspalum oteroi* for forage improvement (Novo, et al., 2016).

Forage legumes are important components of agrobiodiversity, especially in countries where livestock production contributes largely to their GDP. The report produced by FAO (2010) stated that global germplasm holdings had 651,024 forage accessions; 35% were wild species, 13% were landraces, 3% were breeding materials, 4% were advanced cultivars, and 45% were others. 15% of the total genebank accessions collected from 1996 to 2007 were forages and, in total, contribute to 9% of the

major crop groups in total *ex-situ* collections. Australia is the predominant holder of forage legume germplasm, holding 15% of the world's clover holdings. Germany, Japan and Poland have the largest collections of forage grasses. Currently, 74,158 *Trifolium* accessions are maintained in germplasm centres worldwide (FAO, 2010).

The development of information management systems has revolutionised the way germplasm centres store and share data. However, despite the success of data management, the amount of data analysis performed on accessions in genebanks is lacking, and challenges remain to develop useful statistics to strategize wild germplasm use. Graudal, et al. (1995) recognised that the main challenge with conserving plant genetic resources is knowing what to conserve, but accession data can inform decision making (Singh, et al., 2019). As the extent of loss of biodiversity has been made more apparent, the need to efficiently conserve biodiversity is of foremost importance. The increased pressure to secure the genetic diversity of a species over recent decades has created an influx of germplasm accessions into genebanks. With few exceptions, germplasm banks are underfunded and underutilized by breeders (Allard, et al., 1993b).

2.4.1 Core collections

A core collection is a subset of a larger collection (Frankel, et al., 1984), defined as containing a minimum set of germplasm that represents maximum diversity (Zhang, et al., 2019). The development of core collections has seen more urgency in the last two decades (Basigalup, et al., 1995, Johnson, et al., 1999a, Johnson, et al., 1999b, Marita, et al., 2000) and are primarily used as a management tool (Allard, et al., 1993b). The highly characterised accessions within the collection can be used to inform decisions in breeding programmes (Abdi, et al., 2020, Zhang, et al., 2019).

A significant concern of core collections is the potential loss of diversity within the collection. A core collection that does not encompass a considerable amount of the whole collections' diversity would not serve a purpose (Brown, 1989a). To overcome the potential loss of diversity, Brown (1989b) suggested a simple random sampling method which had high retention of diversity statistics. To ensure that core collections are not formed using misleading information, deep characterisation of the germplasm data associated with the accessions compiled in the core collection is needed (Allard, et al., 1993b, Singh, et al., 2019).

2.4.2 Characterisation of germplasm in genebanks

Most of the genetic variation present in genebanks are absent in breeding programmes but will be useful for future breeding programmes (Dwivedi, et al., 2017). Díez, et al. (2018) summarised and identified that the most urgent task in improving genebanks is improving the information available through deep characterisation. Although there is increased genetic and phenotypic data on traits, knowledge of all of the accessions in the genebank as a whole, is lacking (Bretting, 2018). When an accession is collected, it is linked with passport data; geographical information, including latitude and

longitude, and ecological data. Each time the accession is grown in a trial or for regeneration, it is critical that more phenotypic data is collected to characterise the accession more thoroughly in the genebank (Allard, et al., 1991). The information is needed to cater to future agricultural needs (Bretting, 2018).

While the first decades of germplasm centres focussed on conserving genetic variation in species, it has become more apparent that the information about the germplasm has become just as important as the germplasm. Germplasm centres can leverage and exploit data to increase the efficiency and effectiveness of conservation and characterisation efforts of germplasm, as well as increasing the value of the germplasm (Halewood, et al., 2018, Rubenstein, et al., 2006). Studies have already shown that by analysing data in germplasm centres, valuable information can be obtained that can inform future breeding decisions (González, et al., 2018, Ramírez-Villegas, et al., 2010). Egan, et al. (2019a) and Egan, et al. (2019b) characterised the national germplasm collections of red and white clover in the MFGC the national forage genebank in New Zealand. Breeding pools and influential founders were identified, and the authors provided avenues for pre-breeding decisions in the future. Bruce, et al. (2019) demonstrated that although the genetic diversity in Canadian soybean has been maintained for over a century of soybean breeding, the breeders have only used a small portion of the genetic diversity within the available germplasm collections.

The limitation on efficient germplasm exchange is largely based on the information available on the germplasm (Allard, et al., 1993a). As the exchange of germplasm between countries increases, so does the global germplasm network (Allard, et al., 1993a). Annually, the United States shares more than 230,000 seed samples to 100 countries worldwide (Allard, et al., 1991). From 1990-1999, the United States National Plant Germplasm System (NPGS) distributed 621,238 germplasm samples of 10 major crops (Rubenstein, et al., 2006). However, the incompleteness of accession data is a major contribution to the lack of utilisation of germplasm by breeders (Peeters, et al., 1984). Studies have shown that increased characterisation of the germplasm will increase the utilisation of germplasm (Byrne, et al., 2018, Rubenstein, et al., 2006, Yu, et al., 2016).

2.5 Plant breeding avenues for germplasm exploration and retaining and maximising diversity

Plant breeding relies on genetic variance to succeed. Although genetic variance can be well defined, genetic diversity is a term that has no clear definition. It is broadly referred to as any variation at any phenotypic or molecular level in the species at any given time (Fu, 2015). Yet, it is the broadness of the subject that has allowed successful advances, almost exclusively based on useful phenotypes or traits.

Crossing elite germplasm lines to increase the performance of populations has produced cultivars across all major species. Introgression of wild germplasm characters into adapted cultivars is

frequently performed through backcrossing to the elite parent. The development of hybrids through pre-breeding has many challenges, including infertility, linkage drag and crossing incompatibility (Acquaah, 2012). Despite these challenges, pre-breeding, the early activities used to characterise germplasm that identifies useful characteristics (Acquaah, 2012), is frequently used to develop populations with increased variation (Acosta-Gallegos, et al., 2007, Nass, et al., 2000, Sharma, et al., 2013). Identifying genetic variation and utilising information from genebanks to plant breeding programmes is an important strategy for continuing crop genetic improvement (Sehgal, et al., 2015).

2.5.1 Pedigree maps and analysis

Pedigree analysis has been popular in conservation programmes to monitor crosses between individuals to maximise diversity. In the literature, there are more reports of pedigree analysis in animals (Calboli, et al., 2008, Cervantes, et al., 2008, Graczyk, et al., 2015, Hamann, et al., 2008, Leroy, et al., 2006, Roughsedge, et al., 1999, Valera, et al., 2005) than in crops (Gizlice, et al., 1994, Navabi, et al., 2014, Sneller, 1994, Souza, et al., 1989). Although pedigree analysis is typical in animals as both parents are known, it is an important method that enables the characterisation of plant germplasm accessions (Philipp, et al., 2018).

The development of a pedigree map allows a visual representation of population structure and determines the effect of human-based decision making. Pedigree maps show past breeding performed, and breeding pools can be identified (Holland, et al., 2003). It is not crucial that all of the mating relationships are recorded, although indeed helpful (Shaw, et al., 2014). Phenotypic and genotypic information can be added to pedigrees to add depth and knowledge which aids in selection cycle decision making (Breseghello, et al., 2013). However, the large and complex nature of pedigree data sets provides perceptive limitations in building, visualising and analysing large pedigrees (Shaw, et al., 2014). Genetic factors of populations can be described by deriving pedigree-related coefficients such as kinship, inbreeding, the effective number of founders and the effective population size (Bernardo, 2010, Falconer, 1975, Voorrips, et al., 2012). The information obtained from pedigree maps and analysis enables faster and more efficient breeding decisions (Holland, et al., 2003).

Where available, molecular tools are utilised with pedigrees to be used in mixed models (Bink, et al., 2002, Dreisigacker, et al., 2004, Graner, et al., 1994, Janick, 2003, Melchinger, et al., 1994). The integration of molecular markers to validate pedigrees has become increasingly popular to confirm results (Daetwyler, et al., 2012, Paiva, et al., 2011, Smith, et al., 2000, VandeBerg, et al., 1990).

2.5.2 Genome-wide association studies

The ability to describe the relationship between genotype and phenotype has become increasingly important with increased pressure on the performance of crops (Liu, et al., 2018). GWAS is a method that involves scanning the genome using DNA markers to detect associations with phenotypic traits. GWAS was first developed in humans before becoming utilised in plant species. It is a powerful tool

to identify QTL and causative SNPs in both simple and complex traits and can characterise rare variants in species.

GWAS has become popular for analysing simple traits and for furthering the understanding of the genetic architecture of complex traits, i.e. the number of loci that contribute to a trait and the relative contribution to the phenotype. Complex traits are controlled by many rare variants having a sizeable phenotypic effect or, many common variants resulting in a small phenotypic effect (Korte, et al., 2013). Many of the traits that are breeding targets in forages are complex, so GWAS is promising to identify genomic regions controlling traits (Korte, et al., 2013). However, simple traits, traits underpinned by a small number of loci with large effect sizes, are typically best suited for GWAS.

GBS has become a popular sequencing method to use in GWAS studies due to the low cost, high throughput and robustness of the method (Han, et al., 2018, Sakiroglu, et al., 2017, Sonah, et al., 2015). Restriction enzymes are used to reduce genome complexity and the number of repetitive elements. GBS was first developed by Elshire, et al. (2011) and is suitable for outbreeding populations as genome-wide allele frequency profiles can be calculated in pooled samples. However, GBS can pose challenges in the form of low sequencing depth and missing genotype calls (Ashraf, et al., 2016).

The history and successful development of GWAS techniques have been well documented (Ikegawa, 2012, Visscher, et al., 2017). GWAS has been successful in identifying novel variant-trait associations (Tam, et al., 2019) and allowed marker-assisted selection breeding programmes to be developed (Barrett, et al., 2009, Barrett, et al., 2006, Barrett, et al., 2001, Dolstra, et al., 2007, Hayward, et al., 1994, Riday, 2011, Roldán-Ruiz, et al., 2010). However, the number of GWAS studies in forages compared to other crops is low. The highly heterozygous and outcrossing nature of forages makes finding and validating associations more complex compared to other crops that have a closed mating system. Although the overall number of studies is low, there have been numerous GWAS studies in alfalfa that have identified regions of the genome that control forage yield, nutritive value (Sakiroglu, et al., 2012), forage quality traits (Biazzi, et al., 2017), as well as plant growth and forage production under abiotic stresses (Liu, et al., 2017). Significant marker-trait associations have been identified in ryegrass (Arojju, et al., 2016, Brazauskas, et al., 2011, Fè, et al., 2015). There has only been one reported GWAS in white clover where Inostroza, et al. (2018) identified 53 loci associated with cold-tolerance traits.

2.6 Past successes of using germplasm in forage breeding

Utilising germplasm held in genebanks has been crucial for the improvement of plant species, and this has been recognised for many years (Ghimiray, et al., 2017). Hybridising species with germplasm or wild relatives has been successful in plant species. Broadening the breeding pool has introduced increased resistance, yield and variation to be incorporated into populations. Genetic variation is crucial to have in a population, as without variation, genetic gain cannot be realized (Baker, et al.,

1987). Forages have been a group of interest for hybridisation with wild relatives due to the slow rate of genetic gain (Hoyos-Villegas, et al., 2019, Nass, et al., 2012). As perennials, utilization of dominant and additive variance effects could be a feasible avenue to increase the rate of genetic gain in forages. However, this would require modifications to the breeding systems currently in place. Woodfield, et al. (2001a) suggested that the limited rate of genetic gain could be increased by changing from synthetic varieties to hybrid varieties and could provide better control over traits that would improve performance in the species. It is unknown whether environmental adaptation would be sacrificed by such a change, but this could be overcome by the utilization of reciprocal recurrent selection programs based on biparental and test crosses with the deployment of varieties as 'multi-hybrid' composites (vis-a-vis multi-lines) and not synthetics.

The utilisation of species that occur in the primary and secondary gene pool to the target species have been useful to generate new variation to widen the genetic base and improve species (Acquaah, 2012). Ellison, et al. (2006) developed the 'white clover species complex', outlining the species that are closely related and can cross with another in the complex. The most successful example of using germplasm in *Trifolium* is the ISHs of *T. repens* with *T. ambiguum, T. uniflorum* and *T. occidentale* (Hussain, et al., 2016, Marshall, et al., 2008, Marshall, et al., 2015, Nichols, et al., 2014b, Nichols, et al., 2014c, Nichols, et al., 2014d, Nichols, et al., 2014e, Nichols, et al., 2015, Widdup, et al., 2011, Williams, 2014, Williams, et al., 2013, Williams, et al., 2008).

The ploidy of the related species is challenging when developing hybrids. *T. uniflorum* is a tetraploid, *T. occidentale* is a diploid clover, and *T. ambiguum* can exist in the ploidy forms of 2x, 4x and 6x (Abberton, 2007, Jahufer, et al., 2013b, Williams, et al., 2006). The 6*x* form of *T. ambiguum* is the form that is best suited for agronomic conditions, but this has not been able to be successfully crossed to white clover to produce fertile hybrids. Williams, et al. (2008) overcame the genetic bridge of *T. ambiguum* × *T. repens* by doubling the chromosome number and then backcrossing to white clover until stable tetraploid hybrids were produced. More recently, a study by Williams, et al. (2019) used 2*x T. occidentale* × 6*x T. ambiguum* as a genetic bridge to be able to hybridise the two species and produce one gene pool.

The advancements of the ISHs have been successful. Nichols, et al. (2014c) showed that *T. repens* × *T. uniflorum* BC1 hybrids outperform white clover in drought conditions. Nichols, et al. (2014b) identified that some *T. repens* × *T. uniflorum* ISHs that were more tolerant of low external phosphate supply. The next steps for the progression of the ISHs are the continuation of selection cycles to improve populations.

Attempts at hybridisation between other *Trifolium* species have seen some success (Ferguson, et al., 1990). *T. nigrescens* has demonstrated a close affinity with *T. repens* and has had successful crosses (Brewbaker, et al., 1953, Hovin, 1962). *T. nigrescens* has several useful reproductive traits that could benefit *T. repens*, including the prolific number of inflorescences. Marshall, et al. (1995) showed that

T. repens × *T. nigrescens* hybrid progeny showed intermediate reproductive phenotypes and was a significant increase from *T. repens*. More recently, Marshall, et al. (2008) showed that introgression of reproductive traits from *T. nigrescens* to *T. repens* increased the seed yield. Malaviya, et al. (2018) investigated the interspecies incompatibility and affinity between *T. alexandrinum* and 22 *Trifolium* species. Although there was incompatibility among most of the crosses, embryo rescue and intensive crossing produced successful crosses.

Forage grasses have been a target for interspecific hybridisation to improve traits (Aguilera, et al., 2011, Arcioni, et al., 1983, Hunt, et al., 1989, Wilkins, et al., 2003). A prominent early study in New Zealand was the development of the well-known New Zealand short-rotation ryegrass, derived from a cross of *Lolium multiflorum* (Italian ryegrass) × *Lolium perenne* (perennial ryegrass) (Corkill, 1945). The hybrids have been successful in having the rapid and high production from Italian ryegrass and the better persistence and winterhardiness of perennial ryegrass (Arcioni, et al., 1983, van Dijk, 1979). Early work from Takamizo, et al. (1991) demonstrated that Tall fescue (*Festuca arundinacea* Schreb.) and Italian ryegrass (*Lolium multiflorum* Lam.) could be crossed and produce hybrids. The hybrids were the first flowering intergeneric somatic hybrids developed in *Gramineae.* In New Zealand, Widdup, et al. (1992) produced ryegrass hybrids between wild ecotypes and New Zealand or European cultivars. The hybrids showed improved cool season activity and summer quality.

Increased tolerance to drought has been a major breeding target for forage grass hybridisation (Durand, et al., 1997). Perennial ryegrass is a common pastoral grass but suffers under climate extremes. Macleod, et al. (2013) hybridised perennial ryegrass and *Festuca pratensis* (meadow fescue). The resulting hybrids reduced rainfall runoff through the initial large root system that is rapidly produced. The phenotype of the hybrids shows promise for continuing production under changing climate conditions.

Medicago sativa (alfalfa) is a species of interest for hybridisation with related species. Several commercial varieties including Ranger, Vernal and Magnum alfalfa have all been developed through hybridisation. Commercial alfalfa hybrid varieties show large increases in genetic gain compared to non-hybrid varieties (Wiersma, 2001). The first hybrid in alfalfa was developed in 1940. Since then there have been many studies investigating the species that can be introgressed into alfalfa (McCoy, 1985, McCoy, et al., 1986, Nenz, et al., 1996), and produce hybrids that have increased performance (McCoy, et al., 1984).

2.7 Summary

Like many major crops around the world, the future of *Trifolium* breeding in New Zealand will have significant challenges. As the global human population increases, so will the intensification of agriculture. Breeding targets will encompass pest and disease resistance while increasing forage

quality and production. At the same time, there are the looming sustainability goals of countries that aim to have a minimal environmental footprint.

The history and challenges of *Trifolium* breeding can provide context as to why there has been a slow rate of genetic gain. The development of populations, such as *Trifolium* ISHs, that perform in adverse conditions could utilise the nitrogen fixating ability of the *Trifolium* species and reduce farm input. Utilising germplasm will become exponentially valuable to meet the national and global goals of productivity and sustainability. Both the public and private sectors will have to work in harmony to reach these goals, utilising novel and new germplasm and developing innovative methods to adopt into breeding programmes.

Although germplasm centres are crucial to the security of agriculture globally, there has been limited research on the deep characterisation of the accessions held in the genebanks. Pedigree analysis and GWAS are techniques to develop an understanding of the variation present in populations and the combination of both, has increased the power of studies and developed new methods of interrogating and exploiting variation.

Chapter 3

Identification of founding accessions and patterns of relatedness and inbreeding derived from historical pedigree data in a white clover germplasm collection in New Zealand

This chapter has been published in Crop Science:

Egan, L.M., R.W. Hofmann, B.A. Barrett, K. Ghamkhar and V. Hoyos-Villegas. 2019. Identification of Founding Accessions and Patterns of Relatedness and Inbreeding Derived from Historical Pedigree Data in a White Clover Germplasm Collection in New Zealand. Crop Science 59: 2087-2099. doi:10.2135/cropsci2018.11.0688.

3.1 Introduction

White clover (*Trifolium repens* L.) is broadly grown in northwestern Europe and is also grown in association with ryegrass in New Zealand and this combination dominates New Zealand pastures. New Zealand has the highest export share of white clover globally at 57.5% (Rattray, 2005). Approximately 4000 t of 'Grasslands Huia', a New Zealand-bred white clover cultivar, was exported to the United Kingdom in 1980, and 1000 to 1300 t of white clover seed is sown annually in domestic pastures.

White clover has an outcrossing mating system with a gametophytic self-incompatibility system, both leading to high heterozygosity (Abberton, 2007, Williams, et al., 2012). As a highly heterogeneous and diverse species, a platform to monitor the pedigrees is useful (Barrett, et al., 2004, Williams, et al., 2012). Over decades, the aims and goals of white clover breeding have changed due to environmental pressure and market requirements requiring greater physiological and morphological responses to the target environment (Brock, et al., 1989). For example, dry stock farming in New Zealand was in high demand from 1870 to 1920. The improvement in production of yield came mainly from expanding land usage and utilizing soil fertility (Woodfield, et al., 1994), but in the 1990s, New Zealand agriculture introduced pivotal farm cultural practices in conjunction with pasture management to increase pasture performance (Brock, et al., 1989, Woodfield, et al., 1994).

Depending on the objectives of a breeding program, different strategies are adopted to release a cultivar. Many methods have been used in white clover breeding throughout history in New Zealand and worldwide. Mass selection (Wricke, et al., 1986) and recurrent phenotypic selection are examples of common strategies (Caradus, et al., 1998b, Mercer, et al., 2000, Mercer, et al., 1999, Woodfield, et al., 1994, Yamada, et al., 1989). Polycrossing is often used in forage species that display heterosis, where parent clones are grown in isolation and are all pollinated together. The progeny are often combined and tested (Taylor, 2008). However, pollen flow is uneven and the larger the increase in

distance between the male and female genotypes, the lower the chance of successful fertilization (George, 2014). This method is suitable for complex traits if progeny testing is included in the strategy (Taylor, 2008). The outcrossing nature of white clover means that breeding strategies such as mass and recurrent selection use the available variation in the genetic material, while decreasing the risk of inbreeding depression. Although gains continue to occur in white clover breeding (Hoyos-Villegas, et al., 2019, Woodfield, et al., 1994), the genetic consequences of breeding strategies are difficult to assess, particularly if strategies are executed as closed systems with no new genetic variation introduced over long periods of time.

Pedigrees are often used as a conventional method to monitor breeding crosses and population structure in populations of both plants and animals (Navabi, et al., 2014). Pedigree analysis is an important and often essential tool to visualize and describe the population structure and genetic diversity within a population. Molecular tools are being used frequently in mixed models alongside pedigrees (Valera, et al., 2005) in genome-wide association and genomic selection studies (Chen, et al., 2017, Yu, et al., 2006, Zhao, et al., 2011). Although many efforts are focused on conservation breeding program, there is a need for pedigree analysis in breeding programs for plant cultivar development with heavy selection pressure (Jones, et al., 2010). This is particularly the case, if these programs are heavily reliant on germplasm collections. With the advancement in marker and next generation sequencing technologies, there are now methods to analyze relatedness and population structure by using genetic marker data. However, genetic diversity bottlenecks in germplasm stored in gene banks and the capacity and costs associated with genotyping large numbers of individuals or populations may limit the power of these studies. Therefore, information on population structure and relatedness can serve as an appropriate tool in prioritizing plant breeding and genetics efforts. Pedigree analysis has never been performed for a white clover collection in New Zealand or anywhere in the world thus far; the closest source of this kind of information can be found in Caradus, et al. (1997b). The authors in that publication compiled a checklist of white clover cultivars, indicating their parentage and some attributes that relate to the release. However, no relatedness data or quantitative analysis were associated with the handbook.

Safeguarding germplasm is the most inexpensive and efficient method of genetic conservation of wild germplasm of agriculturally valuable plants. The MFGC hosts New Zealand's and international germplasm of forage and pasture plants. The mission of the MFGC is to avail a broad range of genetic diversity in the form of seeds to provide a spectrum of new forage traits to the future breeding programs. The most collected species are the commercially prioritized species for cultivar development by breeders worldwide and specifically in New Zealand such as white clover. Still, the forages of the future are also included in collecting and exchange programs and should receive more attention in the future. The MFGC conserves and occasionally regenerates accessions of wild germplasm, domestic and naturalized germplasm, bred lines, and pre-breeding material. This diverse collection makes MFGC unique among other forage collections around the globe.

We used historical pedigree data from the white clover collection maintained at the MFGC in Palmerston North, New Zealand. The objectives of this study were (i) to create a pedigree map of the collection, (ii) to identify founding accessions and determine the effective number of founders, and (iii) to detect patterns affecting inbreeding and kinship.

3.2 Materials and Methods

3.2.1 Data filtering

The term "accession" is used to refer to any seed material entered into the MFGC with an identification number. The terms "Grasslands cultivar" and "Other cultivar" refer to cultivars that were bred and released by different organizations. "Grasslands cultivars" are all cultivars that were trademarked under the "Grasslands" trademark. "Other cultivar" refers to cultivars bred worldwide by other organizations and are not trademarked as "Grasslands."

To date, the MFGC database holds data for 26,703 accessions of white clover from 40 countries. These accessions were recorded over a timeframe of 75 yr, from 1941 to 2016, using a range of breeding techniques including poly and biparental crossing. Of this total number, 13,687 accessions (Figure 3.1) were used in the construction of the pedigree map using Helium, a software that allows the pedigree visualization of large pedigrees (Shaw, et al., 2014). Accessions were categorized to different subsets based on missing data (1633 accessions), their specific lineage as part of breeding efforts (e.g., seed increases and isolated nodes in the pedigree were not included; Figure 3.1). A total of 12,154 accessions were used for the derivation and analysis of relatedness and kinship parameters. These accessions were selected based on the type of cross they involved (e.g., biparental crosses), whether they had full or half parentage indicated, and whether they had decipherable parental information (Figure 3.1). Polycrosses were excluded from the parameter analysis subset because polycrosses do not fit the allele frequency expectations of biparental crosses; as a result, the assumptions for the calculation of inbreeding and kinship values would be underestimated. The data structure was: accession ID, Parent 1, Parent 2, accession date, seed weight, and country of collection. In total, 2479 accessions had specific collection countries listed.

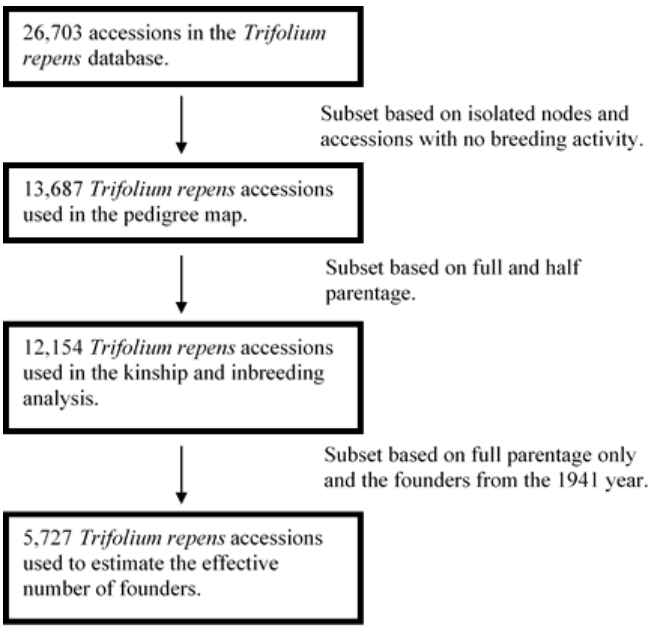

Figure 3.1 Flowchart of the subsetting steps used to filter the *Trifolium repens* accessions used in this study.

Founders were defined as the first accessions in 1941 that had no parentage listed, assuming no breeding had occurred. Every other accession introduced into the database after 1941 was considered as an introduction. Likewise, parents with high contributions to pedigree size were declared as those having families with >50 full or half-sib families. The 50-offspring cutoff was defined arbitrarily as a value well above the mean number of families.

3.2.2 Data analysis

The number of offspring, kinship, and inbreeding was calculated by the R package 'pedigree' (Coster, 2015). Kinship was calculated by using the pedigree information and using a recursive application of the two formulas:

$$F_{yy} = 1/2\left(1 + Fm_y\, f_y\right) \quad [1]$$

$$F_{xy} = 1/2\left(F_{xmy} + F_{xfy}\right) \quad [2]$$

where F is the coefficient of inbreeding, and the kinship of two individuals, given x is not a descendant of y, is F_{xy}. In Eq. [2], $F_{xy} = 0$ when x and y are both from the founder population (Fernando, et al., 2006). The two genes, one from each parent, at a given neutral locus inherited randomly, are transcribed as *my* and *fy* of *y*. To quantify the relationship between *my* and *fy*, the coefficient of kinship is calculated between the two genes. The coefficient of kinship of *y* with itself is Eq. [1]. The kinship coefficient between *x* and *y* is Eq. [2]. Influential parents were defined as accessions with the highest mean kinship (*k*) within their corresponding cluster.

A heat map was used to visualize the relatedness pairwise comparisons of the population among the 12,154 accessions (Figure 3.2a). A dendrogram was used to represent the clustering of the population based on kinship coefficients (Figure 3.2b). Important ancestors were identified as the accessions with the highest mean kinship found at the 22,500 distance coefficient on the dendrogram. This was confirmed by pedigree lineages.

Inbreeding was calculated using the formula (Crow, et al., 1970, Wiggans, et al., 1995, Wright, 1984b):

$$F_y = \frac{H_o - H}{H_o}$$

The unconditional probability that *y* is heterozygous at any given locus is symbolized by *H*. The conditional probability that *y* is heterozygous at a given locus where the genes are not identical-by-descent is symbolized by H_o (Fernando, et al., 2006). Founders were not included in the kinship or inbreeding analysis, as they did not have pedigree data associated with their records. The effective number of founders (f_e) was calculated by the program Pypedal, using this equation from Cole (2007):

$$f_e = \frac{1}{\sum\left(p_i^2\right)}$$

where p_i is the proportion of the genes of the living, descendant population contributed by founder *i* (Lacy, 1989).

The dataset used on the calculation of f_e was the full parental dataset of *T. repens*, containing 5727 accessions (Figure 3.1). Only accessions with full parental data and original founders were used. Attempts to use the full dataset including half parentage did not result in a reliable (inflated) estimate of f_e.

3.3 Results and Discussion

3.3.1 Pedigree map size and complexity

Between 0 and 15 generations were traced. Of the 13,687 accessions, 4724 (34.51%) had full parentage, 7430 accessions (54.29%) had one parent listed, and 1533 accessions (11.20%) had no parentage listed. There were 11,643 terminal lines identified in the pedigree. Terminal lines are defined as lines that are at the end of a lineage. Completeness of parentage (COP) across the entire map was generally high. The COP is defined as how complete the immediate pedigree for the accession was (i.e., one, two, or no parents listed). The COP for each accession was assessed for 16 generations. The mean COP for the reference population was 72%. Generations 5, 13, 14, and 15 had a mean COP of >80%. In general, when the number of accessions per generation decreased, COP increased (Figure 3.3). Within the pedigree map, we identified five notable parents, which contributed

a substantial number of progenies to the overall population. The mean offspring number for the whole population was 5.14, and of the accessions with offspring, the mean was 39.52. The node sizing feature of Helium was used, which is based on the contribution of an accession to the overall population. The five parents with the largest nodes on the pedigree map, indicating the largest amounts of offspring were for C2413 (243 offspring), C6525 (181 offspring), C10850 (305 offspring), C15117 (236 offspring), and C19756 (151 offspring). As parents, these accessions made the largest contributions to the pedigree.

Accession C2413 is an accession that arose from a polycross of C72, C759, C792, C809, and C822 and contributed 243 progeny to the population. The progeny from accession C2413 were used to create a further six generations in the population. Included in the third generation was the accession C6525, contributing 181 progeny, and it arose out of a pairwise cross from C4748 and C4785. Accession C10850 had no listed parentage and contributed 305 offspring to the population. C15117 was an introduction from Spain and contributed 236 progeny to the population. Out of the 236 progeny only C18732 was advanced for selection and produced accession C22640, which in turn produced four terminal accessions: C24250, C22938, C24252, and C24251. Accession C19756 was the progeny of C10576 and produced 151 offspring; 19 of these progeny carried on for further breeding.

Accession C15117 was frequently used as a parent in the early stages of white clover breeding. Originating from Spain, it produced 236 accessions, making it one of the most important parents recorded. The drought-tolerant phenotypes often produced in the Mediterranean countries are exceptionally desirable in white clover, explaining why C15117 was well utilized (Cattivelli, et al., 2008).

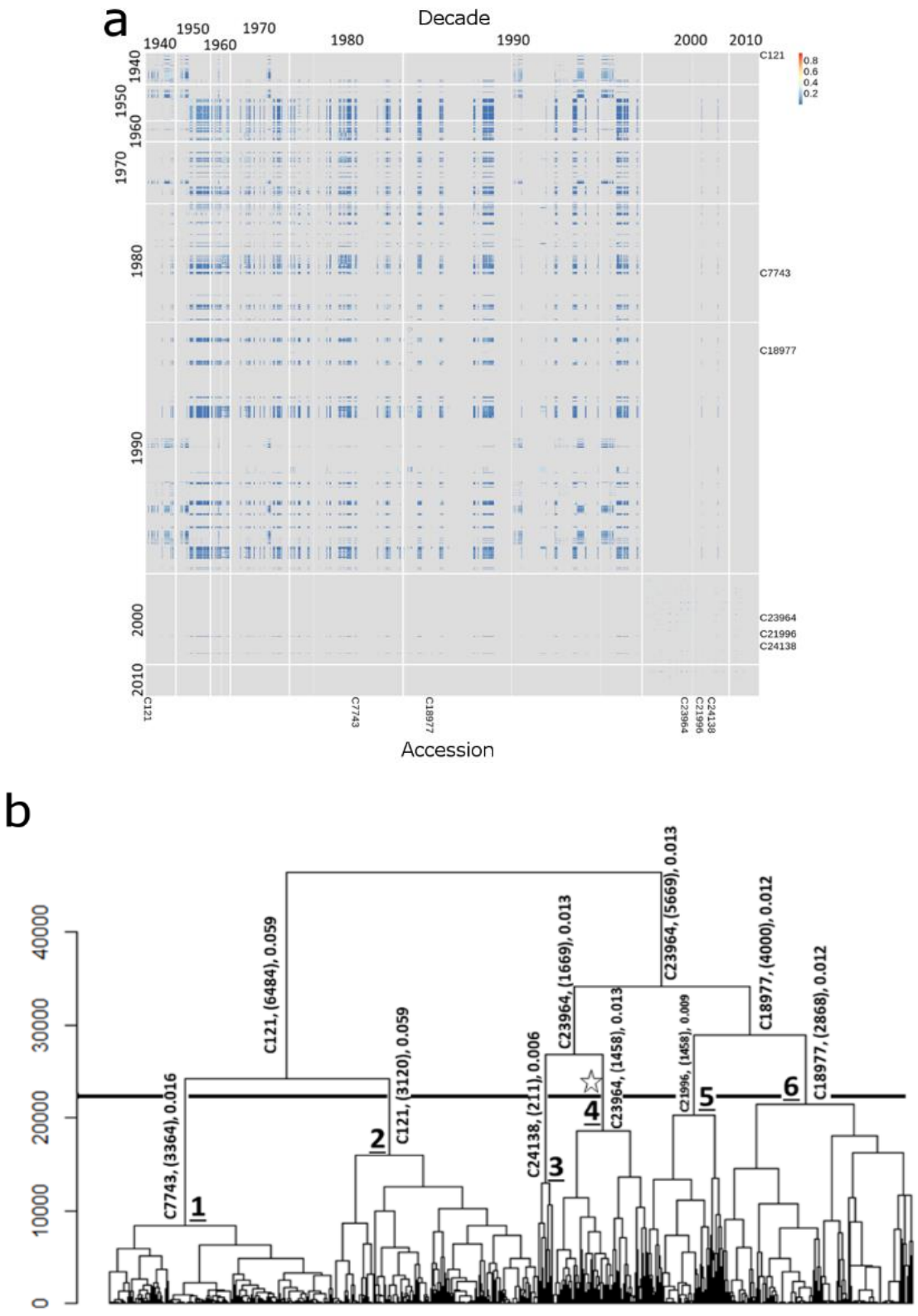

Figure 3.2 The kinship heat map of white clover populations used in this study (a). Dendrogram drawn based on a distance matrix of *Trifolium repens* L. pedigree data (b). Accessions on cluster and sub-cluster are the proposed ancestors and influential parents with the highest average kinship value within a cluster. The solid black line indicates the distance for separating clusters. The number of accessions in each cluster is indicated in parentheses. The star indicates the cluster with commercial cultivars.

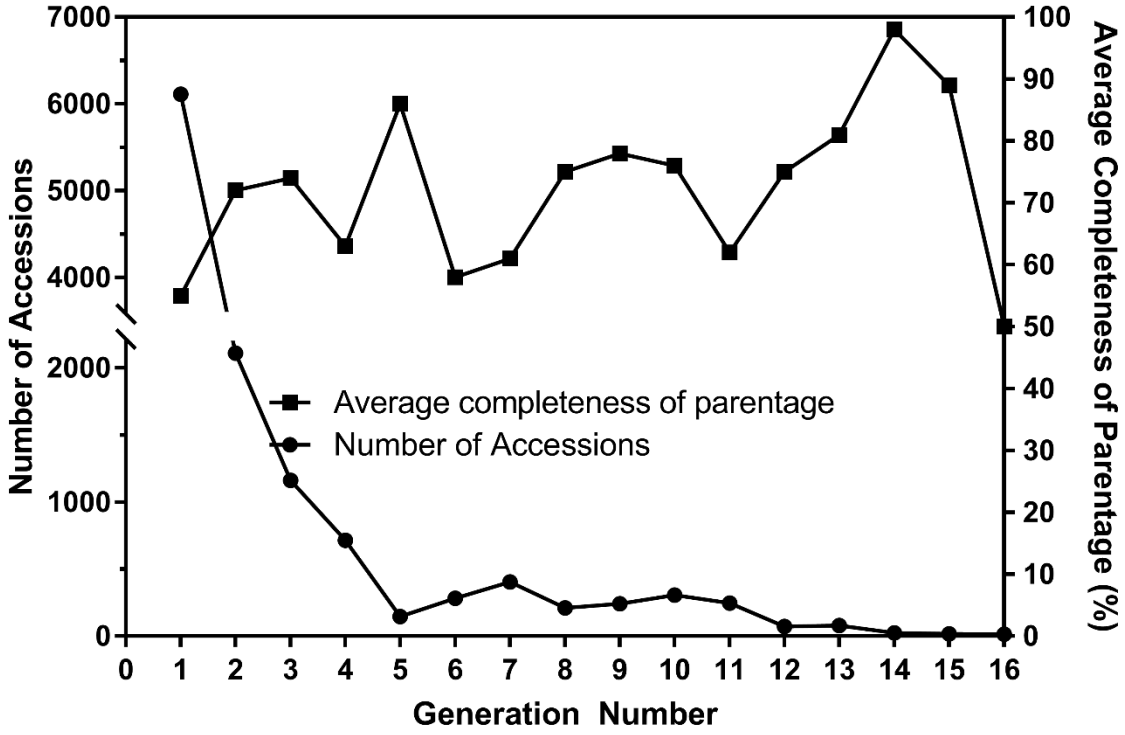

Figure 3.3 White clover collection pedigree data used in this study with the number of accessions and average completeness of parentage per generation across the entire pedigree map.

In the 1940s, there were 22 accessions introduced into the MFGC, all naturalized in New Zealand. The 1950s saw two introductions from Spain and 10 from New Zealand. The 1960s was a decade when the range of geographic origin of introductions rose sharply, indicating the importance of germplasm collecting trips. Accessions were collected from a total of 15 countries, with France, Israel, Morocco, Spain, Lebanon, and Turkey (Figure 3.4) as the most represented countries in the collection.

Figure 3.4 shows the number of white clover introductions into the MFGC for the decades 1950 to 2010. In the 1970s, there was a total of 36 countries contributing to introductions, with France, Germany, Greece, Iran, Israel, Italy, Portugal, Spain, Sweden, Turkey, and the United States contributing to 48.23% of the number of collections. Collections also increased within New Zealand with 182 (37.83%) out of a total of 481 accessions introduced. This pattern carried into the 1980s, where although the number of countries was reduced to 17, the number of accessions was higher, and that included 544 (58.12%) accessions originating from New Zealand. Other countries were Portugal, Spain, and Australia, highlighting the awareness of the requirement to collect diversity from countries with arid environments to enhance adaptation to abiotic stress. In the 1990s, a dramatic increase in collections from the United States, 149 accessions (40.82%) out of a total 365, as well as collections from 27 countries, was recorded. The 2000s and 2010s were representative of a leveling-off period in collecting missions and small increases across several countries, with a noticeable increase in local breeding activity. A total of 294 introductions were collected in the 2000s, and 212 were collected in

the 2010s. The only significantly large introduction in the 2010s was from Russia, where 112 (52.83%) introductions were collected.

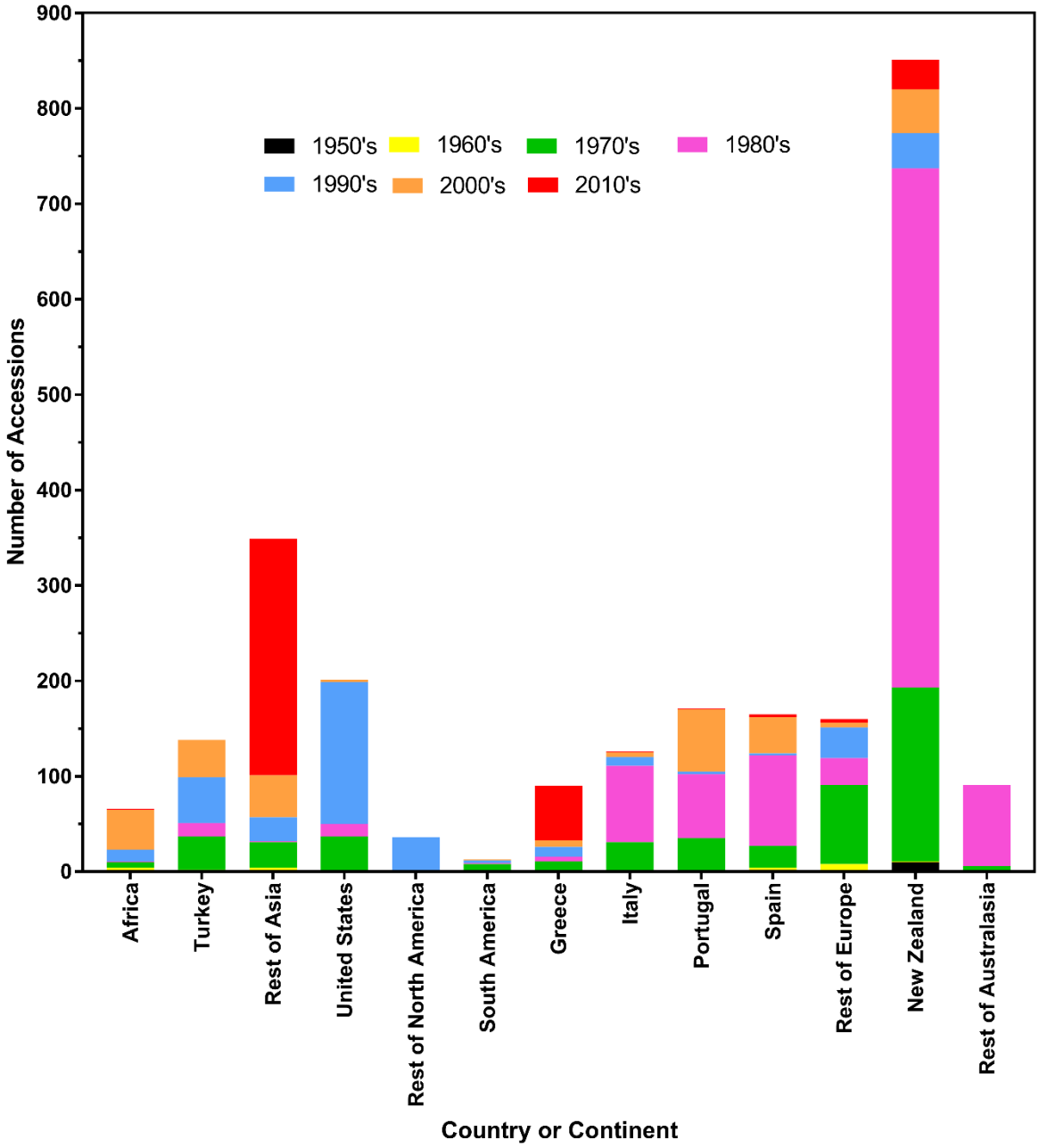

Figure 3.4 The total number of white clover introductions into the Margot Forde Germplasm Centre across the decades used in this study and the geographic origin of the introductions from the decades 1950 to 2010.

The peak in both the New Zealand and international introductions was in the 1980s (Figure 3.5a). From the 1990s onward, fewer accessions were collected from New Zealand compared with international collections. Biparental and single crosses were a common method, whereas polycrossing was rare (Figure 3.5b), with a peak of the practice in the 1990s.

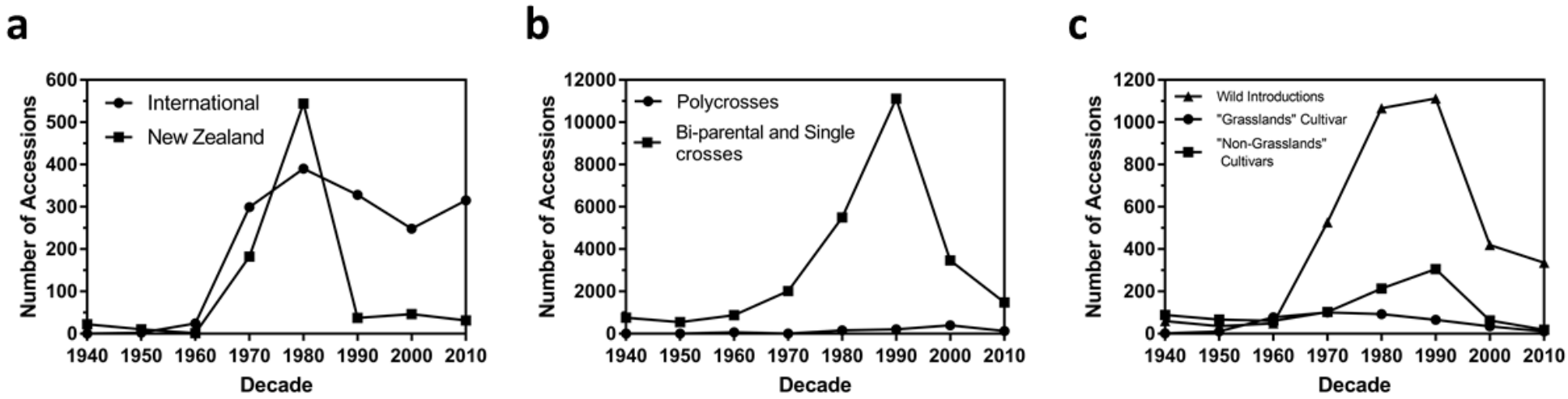

Figure 3.5 (a) The number of international and domestic white clover accessions introduced into the Margot Forde Germplasm Centre per decade between 1940 and 2016. (b) The number of polycrosses and biparental crosses performed per decade. (c) The number of accessions from foreign countries vs. the number of accessions associated with Grasslands cultivars and Non-Grasslands cultivars released per decade.

The idea that there is untapped variation in wild germplasm that can be brought into germplasm centers and breeding programs has motivated the interest in collection trips globally (Hawkes, 1977, Richards, et al., 2010). The numbers of introduced accessions show that wild germplasm has been traditionally recognized as an important source for breeding white clover in New Zealand (Figure 3.4 and 3.5a). Accession C40 was a founding ancestor of the *Trifolium repens* L. section of MFGC classified as "wild" and greatly influenced a large proportion of the structure of the collection.

Many of the influencing ancestors of the population can be traced back to introduced germplasm from various countries, showing the influence of wild germplasm (Figure 3.5c). With increased biosecurity laws, early collection trips provided a wide base of available plant genetic resources. Further expansion of the collection will enable breeding high-performing white clover cultivars to address the future challenges of agriculture.

3.3.2 Offspring distribution

Of the 12.9% of accessions that had recorded progeny, offspring number ranged from 1 to 7,356. The mode of the progeny count distribution was 0, and the mean was 5.14. When accessions with 0 offspring were excluded, the mode was 1 and the mean was 39.52. Inspection of the pedigree maps and offspring distribution suggested that accessions with the highest usage as parents resulted in 90 parents with >50 offspring. Influential parents were introduced across a wide range of years, and no geographical collection data were available for these accessions. The year with the highest number of influential parents (9) was 1941, and the most prominent decade was the 1980s with 26.67% of influential parents, closely followed by the 1970s with 22.22%.

There were 12 accessions that had >1000 offspring, ranging from 1359 to 7356 offspring, with six of the accessions coming from the 1970s. Accessions C72 and C1067 were the two most influential parents in that decade. Accession C72 was referred to as 'North Canterbury type 1', a white clover with a desirable phenotype, and C1067 was an introduction from Spain. C72 had 2860 offspring, and C1067 had 7356 offspring.

3.3.3 Founding ancestors and influential parents

Founders and founder effects

Founders are defined here as accessions that initiated a population (Ladizinsky, 1985). A total of 34 accessions from 1941 were found in the white clover germplasm database as having no recorded parentage and were thus declared as founder accessions.

The overall mean relatedness and inbreeding level of the MFGC white clover collection was <4%. Also, visual inspection of the pedigree map did not result in any obvious bottlenecks. However, confirmation of bottlenecks or founder effects using marker data would be required (Reynolds, et al., 2013). The majority of the introductions were made between the 1970s and 1990s (Figure 3.5a). This

agrees with Figure 3.6, as the kinship and inbreeding levels decrease when the population reaches the 1990s, possibly due to new germplasm integrated into breeding programs.

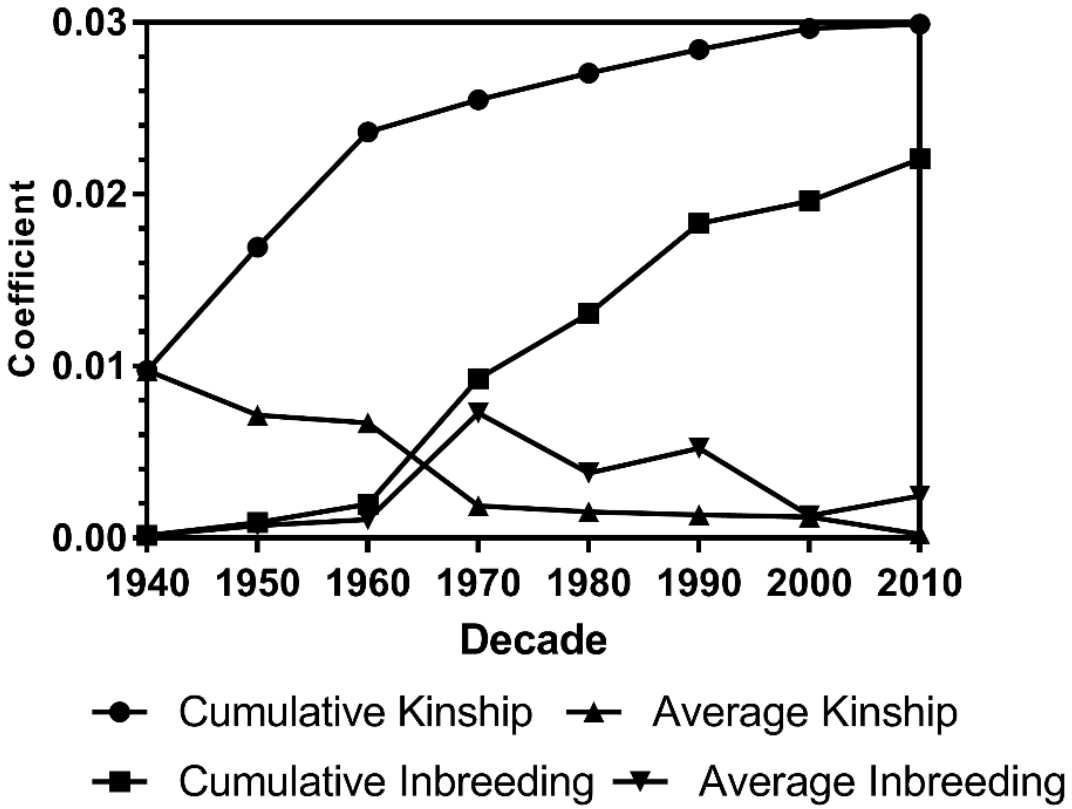

Figure 3.6 The trend in average and cumulative kinship and inbreeding in *Trifolium repens* L. accessions used in this study across eight decades.

The earliest literature contains reports of 'Type 1' white clover, an ecotype that was certified and commercially produced in 1930 (Caradus, et al., 1996c). Three additional ecotypes were also collected, presumably around the same time. Subsequent reselections from these four early ecotypes were recorded under their respective types. When white clover breeding was first recorded, 'Type 1' germplasm performed above the three other types of white clover (Brock, et al., 1989). Most prominently, a breeding program was created to reselect and perform crosses from the commercialized 'Type 1' ecotype population until 1957, where a final selection was completed. This selection was named 'Grasslands Huia' in 1964 and is a current cultivar to date. To clarify, "type" will hereafter refer to accessions belonging to a particular group rather than the original ecotype collected.

The four main types of white clover found in New Zealand pastures were identified among accessions in our study through differences in genetic distance. Type 1 was 'New Zealand Wild White No.1', a medium-leaved, productive perennial. Type 2 was 'New Zealand Wild White No. 2', a small-leaved perennial that was less productive than Type 1. Type 3 was 'Ordinary New Zealand White', a

medium-leaved, nonpersistent clover. Type 4 was ‘Lax early-flowering New Zealand and ordinary European’, a small-leaved, nonpersistent type (Brock, et al., 1989).

Of the 20 founders that were associated with Types 1, 2, 3, or 4 in the database, 14 were associated with ‘Type 1 clover’. Direct parentage in the pedigree indicated that Type 1 was a class to which 70% of the founders belonged. The remaining 14 founders were not associated with any particular type.

Interestingly, 13 of the 34 founders were collected from regions in New Zealand such as Hawkes Bay, Canterbury, Whenuakura, and Southland. Hawkes Bay was the most common collection site, with nine founders associated with the region. Hawkes Bay is a temperate region of New Zealand but is prone to localized and widespread drought. Being a coastal region, extreme weather patterns are common and strong winds often contribute to erosion in paddocks (Mullan, et al., 2005, Thompson, 1987). Hawkes Bay would then have been an environment with high natural selection pressure suitable for finding white clover germplasm with abiotic stress tolerance.

Out of 34, three founders resulted in distinct lineages associated with highly influential parents with high mean *k* values. Interestingly, accession C40 was known as ‘North Canterbury Type 1’ and was a founder of the original breeding efforts from 1941, collected from the Canterbury region in New Zealand. The impact of the elite Type 1 parents was significant, as C40 was the ancestor of the accession that had the largest influence over the whole population. Accession C40 was a founder and produced 51 direct dependent accessions.

The accession C43 was a founder in the white clover breeding population introduced in 1941. It is described as phenotypically similar to Types 1 and 3 white clovers. Accession C43 contributed 17 progeny to the population, as evidenced from its diverse pedigree. The 17 progeny went on to become successful populations themselves, leading to a greater number of indirect progeny.

Accession C63 was known as ‘Imported Kent Type 5’ and was a founder of the original breeding efforts introduced in 1941. It originated from Kent, UK. Kent is one of the warmest parts of Britain; however, it is prone to high winds and as it boards the River Thames and the North Sea to the north, and the Straits of Dover and the English Channel to the south, it can be prone to flooding. These climatic conditions along with the political relationship between New Zealand and England were also a reason why this accession influenced breeding efforts.

Influential parents

Mean kinship values suggested that there were six highly influencing parents in the population structure (Figure 3.2b). The first influential accession was C121 with $k = 0.059$ (Cluster 2 in Figure 3.2b). The parentage shown for C121 is C104/C101. The parentage for C104 is C64/C40, and the parentage for C101 is C63/C40.

Another influencing ancestor was accession C7743 (k = 0.016), which divided the cluster from C121 (Cluster 1 in Figure 3.2b). In the pedigree, the earliest parentage that could be traced back for C7743 was C963/C40. Accession C963 was collected from Spain but harvested in Australia in 1951. The phenotype of a plant traditionally adapted to semiarid environments such as Spain and Australia was beneficial for breeding programs targeting drought-tolerant traits (Cattivelli, et al., 2008).

Accession C23964 (k = 0.013) was an influential parent (Cluster 4 in Figure 3.2b) in that it had two parental cultivars listed, 'Crusader' and 'Kopu II' (Caradus, et al., 1997b). Both of these cultivars are Grasslands cultivars. 'Grasslands Kopu II', formerly known as 'Ranger', was developed from persistent genotypes that were identified in the fourth year of a trial under rotational sheep grazing. Crusader was bred from pair crosses from five 'Crau' genotypes and six genotypes from Syria, which had been identified as having desirable drought tolerance capabilities and high dry matter yield (Woodfield, et al., 2001b).

Accession C18977 (k = 0.012) was an influential accession and had a large pedigree (Cluster 6 in Figure 3.2b). Accession C18977 was listed in 1996 and arose out of C9473/C12186. This accession is in the same pedigree as C121. The founder C40 and another early accession, C72 (another progeny of C40), were in the ancestry of C18977. These two accessions both fall under the 'North Canterbury Type 1' category, emphasizing the role that the Type 1 class had in population structuring.

Accession C21996 (k = 0.009) was the final influencing ancestor found in the dataset (Cluster 5 in Figure 3.2b). The parentage was C19756/C8421, with the earliest recorded ancestors C4145/(C11225/C11248). Accession C4145 was an introduction from Blonei, Poland, in 1978 and only produced one recorded progeny, C8421, which went on to produce 56 progenies.

As can be observed from the size of Cluster 3 in Figure 3.2b, accession C24138 was the least influential parent with k = 0.006; it had the parentage of C19756/C21047.

The main result from finding influencing ancestors (Clusters 1, 2, and 6 in Figure 3.2b) was the impact of accessions associated with the 'Type 1' phenotype. The accessions bred from parents derived from this class influenced the population structure strongly. The founders, although not directly producing large amounts of progeny, produced high-performing offspring that continue to result in many commercial cultivars. Second, downstream of the dendrogram, clusters started to diverge based on geographic origin and plant breeder decision-making patterns. There are some clear distinctions where the geographic origin and the relevant desirable phenotype would influence the structure, such as accessions introduced from the Mediterranean. In contrast, Cluster 4 in Figure 3.2b is represented by germplasm released by a single plant breeder selecting material. Eight cultivars were found in Cluster 4, containing 12% of the total population.

Effective number of founders

There were 5727 accessions analysed, subset based on full parentage plus founder data. A total of 1004 accessions were identified as founders and 4723 accessions were identified as descendants. The estimated f_e was 175.68, indicating the number of founders needed to recreate the population with the same amount of genetic diversity. Figure 3.7 shows the trend in f_e across decades. The f_e value increases proportionately when the number of founders increases. However, it increases at a faster rate when there is a large reduction in the number of founders. In 1970, there were 310 founders and an f_e of 29.28, compared with 1980, when there were 212 founders and an f_e of 64.22. Cumulatively, f_e remained stable between 1940 and 1970, before an increasing rate between 1970 and 2000 and then a decreasing rate in 2010. This may be due to a large number of crosses performed between 1970 and 1990. Large numbers of crosses showed that as the population size increased through breeding, new introductions declined. As a consequence, the influence of founding accessions was reduced as genetic distance increased between founders and offspring multiple generations downstream.

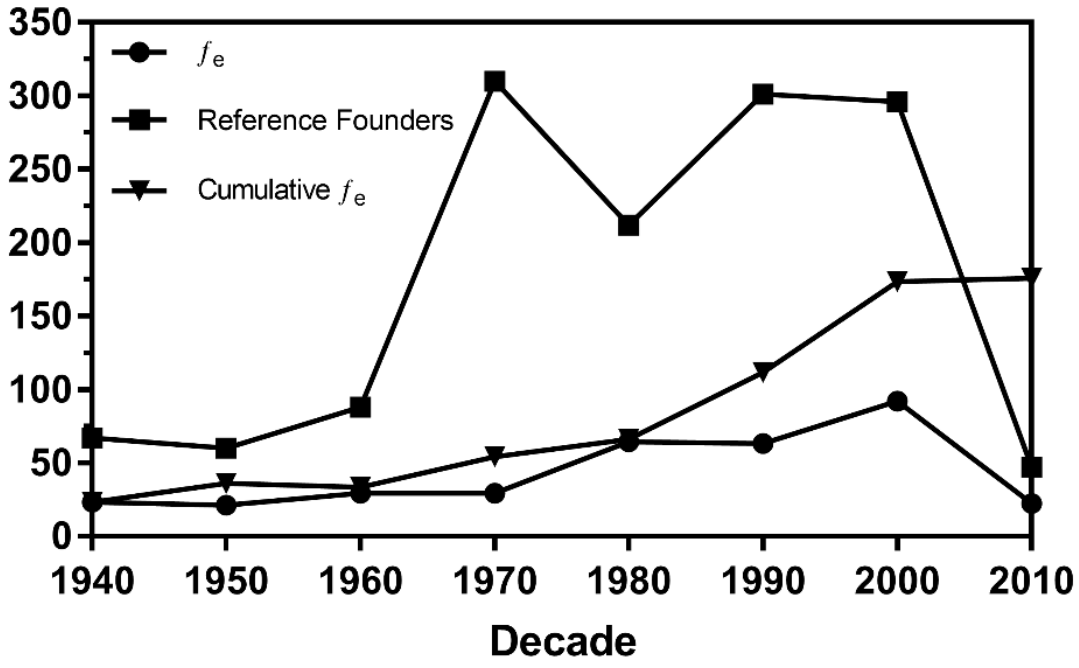

Figure 3.7 The per-decade effective number of founders (f_e) and cumulative across decades in contrast with the reference number of founders with full parentage recorded in the *Trifolium repens* L. collection studied.

Estimates of f_e are useful in predicting future changes in genetic variability (Hamann, et al., 2008). The resulting f_e was 17.50% of the total number of reference founders. A lower number of effective founders relative to the number of reference founders indicates that the contributions of individual reference founders towards the genetic makeup of the population is low. A degree of redundancy in the contributions among founder groups was found, likely between founders from similar genetic background. Some level of redundancy in germplasm collections can serve as a means to minimize

risk of allele loss and increase of inbreeding level in further generations. In practice, our estimate of f_e is also affected by a population with a small group of founders with large contributions.

Interestingly, the effect of an increase in the number of reference founders in a given decade was followed by an increase in f_e one decade later. For example, in the 1970s, there were 310 reference founders with an f_e of 29.28. In the 1980s, there was a decrease in the number of reference founders (212) but a rise in f_e (64.22). This is shown again in the 1990s, when the number of reference founders rose to 301 and there was a steady f_e of 63.17, likely caused by the drop in reference founders in the previous decade. However, in the 2000s, the f_e rose to 92.15 as a consequence of an increase in reference founders in the previous decade. The f_e then stabilized in the 2010s due to no further increases in reference founders in the prior decade.

The impact of founders and introductions on the genetic structure of white clover populations at the MFGC was largely unknown. Without the conformation of a pedigree map, the derivation of inbreeding and kinship coefficients and the visualization of the relatedness in a dendrogram, the population structure and relationships within would be unknown. With this information, the impacts and results of human decision-based breeding over the decades can be evaluated, better information will be available for future planning, and germplasm exchange efforts will be improved.

With increased use of next-generation sequencing, information on population structure remains a key piece of information to guide germplasm surveys and genomic selection efforts. The effectiveness of genomic selection relies on high prediction accuracies. To predict the performance of genomic selection models, it is often useful to simulate prediction accuracies by applying expected prediction accuracy estimators. Often, expected prediction accuracy equations contain a term that requires the number of independent chromosome segments or the effective number of loci (M_e). The M_e term is partially defined by an estimate of the effective population size. An extension of the f_e estimate would allow for an accurate and empirical determination of effective population size (N_e) over a large sample population to be used in genomic selection program planning.

3.3.4 Diversity and inbreeding

Diversity

A total of 73,865,935 pairwise combinations were calculated for kinship coefficient. The values ranged from 0 (no relatedness) to 1 (full relatedness). A heat map was used to visualize the relatedness across the population (Figure 3.2a). The black diagonal represents the perfect relationship of each accession with itself, and the symmetric of diagonal elements represent kinship measures for pairs of accessions. Relatedness is indicated by color intensity. Overall mean relatedness was $k = 0.002$. The yellow or red clusters in Figure 3.2a represent high relatedness clusters—for example, C1061/C480 ($k = 0.75$), C1042/C703 ($k = 0.562$), C2146/C1863 ($k = 0.531$), and C104/C64 ($k = 0.5$).

A total of 5529 accessions (45.49%) had a mean kinship of 0, whereas 6625 accessions (54.51%) had a kinship level of <0.2, 96.5% had 0 relatedness, 3.49% had indirect relationships, and 0.01% had half or more kinship. These kinship values indicate that genetic relatedness within the germplasm collection is low.

A total of 96.5% of the pairwise combinations had 0 relatedness, and when all possible combinations were mean per accession, 45.49% showed kinship levels of 0. Accessions that had a mean kinship of 0 in both datasets may show promise for the exploration of divergent parental combinations.

In *T. repens*, inbreeding depression has been proven to affect some morphological traits. Michaelson-Yeates, et al. (1997) used inbred lines of white clover, utilizing the self-fertility (*Sf*) allele. It was noted that only half of the hybrids showed positive heterosis, and no other trait showed significant heterosis. However, the degree of heterosis was related to the extent of variation in morphological characters between the parental lines. Nichols, et al. (2007b) assessed the impact of inbreeding depression on nodal root system morphology. The roots became shorter and thicker, but the root architecture was largely unaffected; however, there was reduced nutrient uptake efficiency compared to the parent clover. These studies emphasize the risk that is associated with inbreeding depression in white clover.

Although both of the studies above show that in white clover inbreeding is deleterious to some morphological traits, it should be acknowledged that when monitored and used correctly, inbreeding can lead to increased genetic gain. Inbreeding can unmask positive recessive genetic variation and can be used to remove unfavorable genetic load (Humphreys, 1997, Rotili, 1991). Hybrid vigor has been shown in crosses between inbred lines (Michaelson-Yeates, et al., 1997). Atwood (1945) used inbred lines produced by the self-fertility allele, and positive heterosis for dry matter production was seen in half of the hybrids.

Accession membership to Clusters 1, 2, and 6 can be largely explained by country of origin in Figure 3.4. Cluster 1 had exclusively accessions from New Zealand (22). Cluster 2 had the largest diversity of countries: Turkey (16 accessions), Portugal (2 accessions), Italy (2 accessions), Spain (2 accessions), Germany (1 accession), Denmark (1 accession), Sweden (1 accession), and Poland (1 accession). Cluster 3 had 28 accessions from Turkey, 9 from Spain, 10 from Portugal, 30 from Georgia, and 15 from Armenia.

Inbreeding

Inspection of the overall pedigree suggested no visible bottlenecks occurring, which relates to the low inbreeding coefficients found. The mean inbreeding level was 0.39%, and among the accessions with nonzero inbreeding coefficient, the mean was 8.83%. The frequency of inbreeding within the inbred accessions peaked at both coefficients 0.04 and 0.13. Among the accessions with nonzero inbreeding coefficient, the highest level of inbreeding was 0.33 and the lowest was 0.0002. Across the whole

dataset, 11,624 accessions (95.64%) had an inbreeding coefficient of 0, whereas 530 accessions (4.36%) showed inbreeding.

The trend in inbreeding from 1940 to 2010 shows that from 1940 to 1960, there was an initial increase in inbreeding of 0.0018. From 1970 to 1990, the steepest increase in inbreeding was found at 0.009. There was an increase from 1990 to 2000 of 0.0013, before another increase of 0.0024 from 2000 to 2010 (Figure 3.6).

The genetic consequences of inbreeding in outcrossing species can be adverse. Due to the high levels of heterozygosity, most outcrossing species carry a high genetic load of deleterious alleles and suffer from severe inbreeding depression (Jones, et al., 2010). Inbreeding depression is the reduced fitness of a population as a result of inbreeding, where the recessive alleles increase in frequency but are less favorable than the dominant alleles, resulting in a reduction in performance.

Inbreeding depression has severe effects in alfalfa (*Medicago sativa* L.). Wilsie (1958) showed that one generation of selfing resulted in a mean loss of 20 to 30% in vegetative vigor and 80 to 90% in self-fertility. Dessureaux, et al. (1969) investigated the pattern of inbreeding depression in two specific alfalfa genotypes and the impact on the first-generation hybrid as the parents became more inbred. Inbreeding depression increased in each generation, and by the third generation, the progeny had practically become self-sterile. In contrast, Busbice, et al. (1966) observed a 30% reduction in forage yield in alfalfa.

Acquaah (2012) indicated that mating systems such as half-sib mating, full-sib mating, and backcrossing can increase inbreeding. Autopolyploids have multiple alleles and can accumulate deleterious alleles that may not show up until later generations. Inbreeding depression is usually more severe in autopolyploids than in diploids; however, the rate to homozygosity is much slower in autopolyploids.

3.3.5 Commercial cultivar development

The dendrogram in Figure 3.2b shows six significant clusters and 10 breeding pools occurring in the population. To assess the impact of the MFGC white clover collection in the development of commercial cultivars of interest to research and industry, accessions associated to the title of "Grasslands cultivar" or "Non-Grasslands cultivar" were extracted from the relatedness data.

Cluster 4 in Figure 3.2b had 37 accessions linked to commercial cultivars. The accessions in Cluster 4 were introduced between July 1996 and January 2016. The geographic origin of commercial accessions within Cluster 4 was confined to Australasia, with agronomic traits common to Australia and New Zealand that influenced the divergence of clusters. Eight accessions (C26366, C26367, C26368, C26594, C26794, C27071, C27072, and C27073) in Cluster 4 had specific commercial cultivars listed in their data. These cultivars were all listed within 4 years of each other, and common

phenotypic characteristics between these cultivars were found (Table 3.1). Six out of the eight cultivars were medium leaved, and another six were also bred with persistence as a breeding objective (Beuselinck, et al., 1994).

Table 3.1 Table of the eight commercial cultivars clustered in cluster 4 in Figure 3.2b and their key attributes (GrasslanzTechnology).

Cultivar	Key attributes
Grasslands™ Legacy	Large leaved, high growth and persistence.
Grasslands™ Tribute	Medium to large leaved, tolerant to clover root weevil, high stolon density:leaf size ratio.
Grasslands™ Nomad	Small to medium leaved, high persistence.
Grasslands™ Bounty	Medium leaved, high yield and persistence, and high stolon density.
Quest	Medium to large leaved, high tolerance to clover root weevil, frost tolerance.
Durana	Medium leaved, high persistence, low-growing.
Avalon	Small to medium leaved, high stolon density.
Grasslands™ Patriot	Medium to large leaved, high persistence and yield.

The majority of New Zealand white clover cultivar releases thus far have occurred between the years of 1970 and 1990, and the increase in cultivar release in these decades is evidenced by the contributions of the number of introductions over the same decades, as shown in Figure 3.5c. Patterns of introductions and releases peaked in 1990 and decreased thereafter. The historical patterns shown in Figure 3.5c are evidence of the direct relationship associated with the role that germplasm centers and collections play in the development of elite cultivars for farmers.

The biological and economic importance of white clover germplasm and breeding to the New Zealand pastoral sector is immense, with high recognition during the 1990s. Mather, et al. (1996) reported that the 1994 Organisation for Economic Cooperation and Development (OECD) Register listed 93 white clover cultivars, with a further 25 to 30 cultivars also known to commerce. Therefore, there is well in excess of 100 cultivars to fill the global annual market of 8500 to 10,500 Tg. As New Zealand provides 50 to 55% of this seed, there is increased motivation and economic benefits to breed elite cultivars for the market, and germplasm forms one the foundations for cultivar development, food, and agriculture (Ghimiray, et al., 2017).

To have the capacity to address climate change, pre-breeding efforts will increasingly need to rely on germplasm banks. Worldwide, there are now >700 seed collections holding an estimated 2.5 million entries (Plucknett, et al., 1987, Tanksley, et al., 1997). For example, the USDA-ARS NPGS is a major source for plant genetic resources worldwide. As of November 2016, the number of accessions in the

USDA-ARS NPGS was 576,325, representing 15,116 species, and in 2015, 239,118 of those accessions were distributed (Byrne, et al., 2018, Clark, et al., 1997). The United States spends approximately US$20 million annually on the maintenance of those collections (Tanksley, et al., 1997).

3.3.6 Effect of forage breeding strategies

Frequent monitoring of breeding programs and breeding decisions prevents the creation of genetic bottlenecks, which limit the ability to generate new genetic variation that enables continued genetic gain. Our results show that albeit low, inbreeding should be paid careful attention going forward. Casler, et al. (1996) and Casler (1998) noted that the lack of improvement in forages despite immense breeding efforts can be attributed primarily to the long breeding cycle, as the majority are perennials, and secondly to the negative correlation between yield and many other economically important traits in the forages (Casler, et al., 2008).

White clover breeding programs have mainly relied on recurrent phenotypic selection and crossing the most elite plants (Bell, 1977, Hill, 2014). Quantitative genetics principles applied to plant breeding developed in the 1940s led to the present time integration of population genetics theory into plant breeding programs, which led to a better understanding of the genetic architecture of traits via next-generation sequencing (Hill, 2014). However, in order for genome-wide association studies to be successful, germplasm and knowledge of the population structure is crucial.

Forage breeding around the world is mainly performed by half-sib family selection. Half-sib family selection reduces genetic gain by half because there is only control over the female parent. Gain can be doubled by selfing each parent to obtain S_1, then crossing to obtain half-sibs (Acquaah, 2012, Wilkins, et al., 2003). These are most often crossed through a polycross mating system and are useful when selecting traits of high heritability. In contrast, full-sib family selection can be generated from biparental crosses using parents from the base population. The families are evaluated and the elite families are selected. The half-sib/full-sib family selection has a number of merits; it has been in place for a long period of time and produced numbers of cultivars, also it is cost effective and resource efficient (Acquaah, 2012). Although straightforward and cost effective, it does not capture the full variation and potential in the population. Reciprocal half-sib selection, also known as reciprocal recurrent selection, is a strategy for interpopulation improvement; two diverging populations are used, and each population is used as a tester to evaluate the other. Reciprocal full-sib selection is used for interpopulation improvement for species where the commercial product is hybrid seed. The selection cycle is completed in the fewest number of seasons by using plants from which both selfed and hybrid seed can be obtained (Fehr, 1991).

Casler, et al. (2008) and Hoyos-Villegas, et al. (2018) proposed theoretical gains that could be captured if among- and within-family selection was used in the forages. Their findings showed that

among- and within-half-sib-family selection is the more efficient and than family selection under all circumstances for any positive value of within family heritability. Among- and within-family selection and progeny testing are more expensive and resource intensive than half-sib/full-sib family selection. However, the theoretical gain that could come from using these techniques is strong enough that the resources used in among- and within-family selection could be justified.

There were significant clusters that diverged from each other in our study (Figure 3.2b), most likely due to different selection pressures. There is both theoretical and empirical evidence that supports the idea that hybrids developed by crossing populations that have diverged can outyield the better performing parental population (Brummer, 1999). However, it is likely that without proper intrapopulation selection, nonfavorable alleles of strong effects will be present, and any benefits from additive × additive variation will not be observed. In contrast, with the right strategy, pre-breeding efforts aimed at generating new genetic variation in white clover will benefit from large-scale population structure information.

3.4 Conclusions

To our knowledge, this is the first study of its kind in white clover. The construction of pedigree maps and relevant demographic information showed that Australia, France, Germany, Greece, Italy, and Spain were the countries that had the most consistent introductions over the 75 yr that the MFGC has collected germplasm. The genetic diversity in white clover was >96%, reflected by low inbreeding levels. Although there was a steep increase in inbreeding from 1970 to 1990, it should be noted that inbreeding did not exceed 2%. Low inbreeding is a positive sign of the amount of diversity contained within the collection, slowing the loss of unique and favorable alleles occurring in the species if properly used.

The founder accessions related to ‘Type 1’ white clover had a large influence on the population. Identification of founder accessions will inform future studies on the uniqueness of germplasm stored at the MFGC in relation to other collections worldwide.

The results of our study allowed the visualization of historical patterns of relatedness and inbreeding in white clover germplasm. The ultimate aim of this process will be to increase genetic gain in white clover. Information obtained from population structure and breeding pools will enable opportunities to perform new crosses, design new breeding strategies among and within clusters, and contribute to better germplasm utilization.

Increases in knowledge and application of quantitative and population genetics models in combination with new technologies and use of interpopulation and intrapopulation pre-breeding strategies will enable efficiencies in breeding. However, germplasm collection remains a critical component to maintain progress.

The limitations of pedigree analysis are largely due to the quality of the records maintained. However, in white clover, a group of plants are often polycrossed. Current pedigree analysis software can only handle two parents, excluding polycross data from analyses, and limits the scope of results.

Overall, estimates of kinship and inbreeding indicate that the MFGC has been successful in maintaining and elevating genetic diversity in white clover. This achievement has been realized by continuous germplasm collection trips and exchange in a demonstrated relationship with cultivar development. Despite successful breeding efforts, the increasing demand for adaptation to climate change and more sustainable animal production requires better and more aggressive utilization of white clover germplasm in the future.

Chapter 4

Identification of founding accessions and patterns of relatedness and inbreeding derived from historical pedigree data in a red clover germplasm collection in New Zealand

This chapter has been published in Crop Science:

Egan, L.M., R.W. Hofmann, K. Ghamkhar and V. Hoyos-Villegas. 2019. Identification of Founding Accessions and Patterns of Relatedness and Inbreeding Derived from Historical Pedigree Data in a Red Clover Germplasm Collection in New Zealand. Crop Science 59: 2100-2108. doi:10.2135/cropsci2019.01.0045.

4.1 Introduction

Red clover (*Trifolium pratense* L.) is a native species in Europe, Western Asia and northwestern Africa. It is also grown widely as a fodder crop and is used as silage and hay. In a grazing system, it is often mixed with white clover (*Trifolium repens* L.) in pasture mixes (Cassileth, 2010, Kemp, et al., 1999). Worldwide, red clover occupies approximately 4 million ha (Riday, 2010) and is an important component of pastures that sustain productivity and income for subsistence farming communities, such as the Aymara in the Andean plateau (López, et al., 1998). Red clover is a key component in forage systems for its N fixation ability (Vleugels, 2013). Carlsson, et al. (2003) reported that the extent of N fixation by red clover can be up to 373 kg N ha^{-1}. Red clover can contain up to 1% estrogenic compounds. Estrogen can interfere with ewe fertility (Morley, et al., 1964) and formononetin is considered the main compound responsible for this condition. Therefore, newly bred cultivars have been developed with low (0.8% formononetin) levels of estrogen compounds (Cassileth, 2010, Kelly, et al., 1980). However, Sutherland, et al. (1980) showed that estrogen can improve red clover's tolerance to pests, such as grass grub [*Costelytra zealandica* (White)] and black beetle [*Heteronychus arator* (Fabricius)].There is some mixed evidence in the literature suggesting that isoflavonoid phytoestrogens have some human health benefits in lowering the risk of osteoporosis, heart disease, and breast cancer (Patisaul, et al., 2010).

Red clover is relatively drought tolerant (Vaseva, et al., 2011) and provides a high-quality feed throughout summer, whereas other species are adversely affected by water deficit. When grazed infrequently during spring and summer it can produce ~12 t dry matter ha^{-1} (Kemp, et al., 1999). Red clover thrives in pastures via rapid establishment and tolerates poorly drained soils (Riday, 2010). Unlike white clover, red clover cannot tolerate hard, continuous grazing (Kemp, et al., 1999) and is mostly used as cut feed. Red clover normally persists for 2 to 4 years but can persist for 7 years where favorable conditions are present (Kemp, et al., 1999, Ledgard, et al., 1990). The red clover cultivars

‘Tuscan’ and ‘Relish’ have been bred for persistency, productivity, and acceptable seed yields (Charlton, et al., 1999). Despite the extent and success of red clover cultivar releases, relatively small gains have been achieved in increasing the persistence of red clover through breeding programs. In a review of breeding for improved persistence in red clover in Chile, Ortega, et al. (2014) reported a mean realized annual genetic gain of 0.4 to 0.8% for red clover persistence. One cultivar, Carillanca, under irrigation had a large annual gain of 2.6%.

Red clover is a diploid ($2n = 2x = 14$) with a genome size of approximately 420 Mb (Sato, et al., 2005). Red clover is almost fully self-sterile but highly fertile when outcrossed (Williams, et al., 1947). The high level of heterozygosity is attributable to the gametophytic self-incompatibility system present (Taylor, 1982). Because of the economic and agricultural importance of red clover, the number of genetic studies has increased in recent decades (De Vega, et al., 2015, Ulloa, et al., 2003). Abberton (2007) reviewed interspecific hybridization between red clover and its related species. These programs had been running for >50 years, usually utilizing embryo rescue techniques. The aim of these programs was to increase persistence in the sward through crosses with wild germplasm, such as *T. medium* L. Although no new cultivars have arisen through this technique, it has expanded knowledge on the evolutionary aspects of the genus.

Red clover cultivars can be either diploid or tetraploid. Tetraploid red clovers ($2n = 28$) can outperform diploids in several aspects (Taylor, 2008), such as dry matter yield. Tetraploids are often produced by chromosome doubling of diploid lines, as described by Taylor, et al. (1996a). Tetraploid plants are generally larger with larger leaves and florets and have increased persistence compared with diploid plants. However, because of the large floret size, pollination can become an issue. Seed yields are therefore lower than in diploids, and this is a limiting factor to the use of tetraploid cultivars (Charlton, et al., 1999, Taylor, et al., 1996a).

Red clover cultivars are described by ploidy level and flowering time. Early-flowering cultivars can be grown in a wider range of environmental conditions and give more frequent but lower yielding cuts than late-flowering cultivars (Abberton, et al., 2011).

The use of pedigrees in plant breeding is a traditional method that has proven successful across many decades (Navabi, et al., 2014). A pedigree map is a visual representation of the relationships within the population, showing relatedness between individuals or groups of individuals (Acquaah, 2012). In plant breeding programs, they are useful for deciphering the population structure and visualizing existing or potential genetic bottlenecks. Phenotypic data are often projected onto the pedigree to visualize the flow of targeted traits throughout the pedigree (Shaw, et al., 2016). Pedigree assembly and analysis also allow for the determination of inbreeding and kinship coefficients in a population, which are useful in assessing the status of genetic diversity in a germplasm collection. When coupled with historical information, co-ancestry analysis can prove a powerful tool in examining the patterns that lead to population differentiation. The inbreeding coefficient is defined as the probability of

drawing two homologous alleles from an individual that are identical by descent, indicating that they result from one allele from a common ancestor. The kinship coefficient is the relatedness between individuals, indicating the proportion of alleles identical by descent shared among individuals (Hedrick, 2011).

Often, when gain for agronomic traits is limited by low genetic variation, the most appropriate way to introduce new variation is to find new germplasm from the primary gene pool. Plant introductions may be of more value in the early stages of breeding programs, rather than using released red clover cultivars as parents (Taylor, et al., 1996a). A successful example is the pedigree of 'Cherokee', a cultivar that was developed for an area of the United States where red clover had not previously been grown. The base population was made up of 75% selections from plant introductions and 25% from two older cultivars. After five selection cycles, a population was produced that had reduced dormancy and high levels of root-knot nematode (*Meloidogyne* spp.) resistance (Quesenberry, et al., 1993).

The increased use of next-generation sequencing technologies has allowed plant breeding programs to accelerate their rates of genetic gain. However, in order for germplasm collections to contribute to better understanding of genetic diversity and the contained population structure, it would be highly valuable to have a detailed pedigree map (Dias, et al., 2008, Kouamé, et al., 1993, Mosjidis, et al., 2006, Taylor, et al., 1996a, van Berloo, et al., 2005). Plant breeding programs use pedigree data to obtain insight into the germplasm, leading to better-informed breeding decisions (van Berloo, et al., 2005). Pedigree assembly can provide insights into genetic bottlenecks, inbreeding depression, and low allele complementarity.

The MFGC is located at the Grasslands campus in Palmerston North, New Zealand. It collects and maintains germplasm of approximately 2000 forage species, including wild and domestic germplasm, and released cultivars, including intraspecific hybrid cultivars.

By providing information resulting from stored pedigrees at the MFGC for pre-breeding efforts, new avenues can be followed to address the limitations of current red clover cultivars (i.e., low persistence or high estrogen levels). To the best of our knowledge, there has been no pedigree analysis of the red clover collection in the MFGC to date.

We used historical pedigree data from the red clover collection maintained at the MFGC. The objectives of this study were (i) to create a pedigree map of the MFGC red clover collection, (ii) to identify founding accessions, and (iii) to detect patterns affecting inbreeding and kinship.

4.2 Materials and Methods

4.2.1 Data filtering

Pedigree map construction was undertaken using the MFGC database, comprising 5223 accessions from 41 countries worldwide. Accessions were curated across a timeframe of 82 years, from 1934 to

2016. Some of these accessions were developed through the application of a range of techniques, including poly- and biparental crossing. Of the total, 3291 accessions were used in the construction of the pedigree maps by Helium, a software program that visualizes large-scale plant pedigrees (Shaw, et al., 2014). Subsetting was performed on the basis of missing data and whether the accessions had continued breeding (i.e., single nodes in the pedigree were removed, as they were indicative of no breeding activity, Figure 4.1).

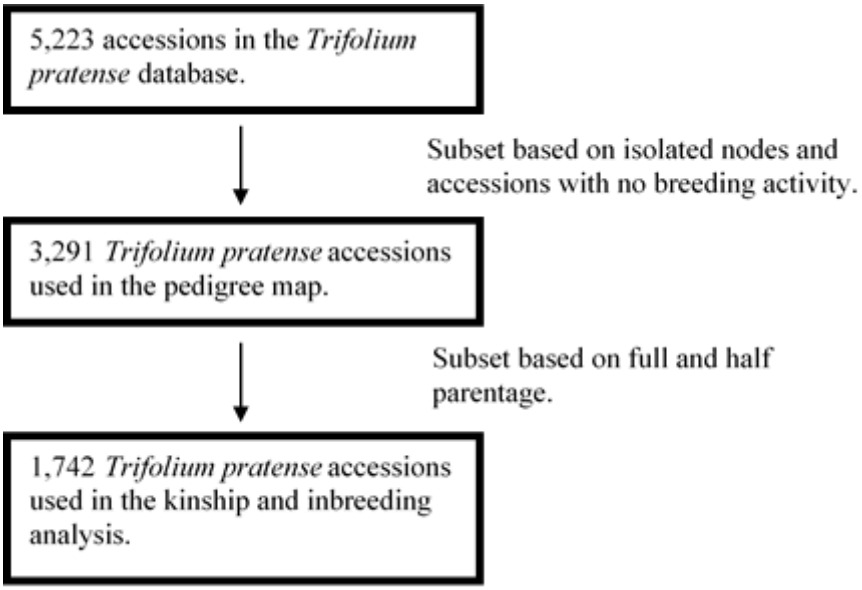

Figure 4.1 Flowchart of the steps undertaken to subset and filter the *Trifolium pratense* accessions used in this study.

Of the total number of accessions, 1742 accessions were used for the derivation and analysis of parameters. Pedigrees associated with accessions were subset based on biparental crosses, with full or half parentage recorded. For clarity, we will refer to half- or full-sib families as "families" throughout the paper, unless specifically stated half- or full-sib. Polycrosses were excluded from the parameter analysis subset because polycrosses do not fit the allele frequency expectations of biparental crosses, and inbreeding and kinship values would have been overestimated.

An extensive search and testing of software packages did not result in an option that could handle (i) the combinations of pedigrees with both poly- and biparental crosses and (ii) the capacity to handle the number of accessions to be analysed. The data structure was accession ID, Parent 1, Parent 2, accession date, and country of origin. Accession ID was the unique number given to the accession when it was entered into the database. Parent 1 and Parent 2 represent accession IDs that were used as parents in the cross. All parents would have been entered into the database earlier and have a unique accession ID. Accession date is the date when the accession was entered into the database. In total, only 2328 accessions had the specific location of country of origin listed. Founders were defined as the accessions from 1934 to 1937 that had no parentage recorded, suggesting no prior breeding. Other accessions introduced into the database after that were considered introductions. Accessions with a

total of more than 80 half- or full-sib families were regarded as influential accessions with significant footprint across the pedigree map. Accessions were clustered by the decade of entry into the MFGC database. It should be noted that the use of the term "accession" refers to any entry previously or currently stored in the MFGC. The term "Grasslands cultivar" represents accessions stored in the MFGC that were submitted as bred and released cultivars under the Grasslands trademark. The term "other cultivar" refers to cultivars bred in New Zealand or internationally by other organizations and not under the Grasslands trademark.

4.2.2 Data analysis

The R package "pedigree" was used to calculate the number of offspring families, kinship, and inbreeding (Coster, 2015). The two formulas below were used in a recursive application to calculate kinship:

$$F_{yy} = 1/2\left(1 + Fm_y\, f_y\right) \quad [1]$$

$$F_{xy} = 1/2\left(F_{xmy} + F_{xfy}\right) \quad [2]$$

where F is the coefficient of inbreeding, x is not a descendant of y, and the kinship of two individuals is F_{xy}. Assuming $F_{xy} = 0$ when x and y are both from the founder population, m_y and f_y of y describe the genes from each parent that are randomly inherited, and the relationship between m_y and f_y is described by the calculated kinship coefficient between m_y and f_y. The kinship coefficient of y with y is Eq. [1]. The kinship coefficient between x and y is Eq. [2] (Fernando, et al., 2006). A heat map was used to visualize the relatedness between 1742 pairwise comparisons of the population (Figure 4.3a). The black diagonal represents perfect relationship of each accession with itself and the symmetric of diagonal elements represents kinship measures for pairs of accessions. A dendrogram was used to represent the clustering of the population based on kinship coefficients (Figure 4.3b). To identify important ancestors in a particular sector of the germplasm, accessions with the highest mean kinship found at the 2000 distance coefficient were identified (solid black line in Figure 4.3b), which resulted in the second or third branch from the main branch on the dendrogram. Further insight into the history of different accessions was verified by studying individual pedigree lineages and the lineage contribution to the population in Helium.

Inbreeding (F_y) was calculated using the formula below (Crow, et al., 1970, Wiggans, et al., 1995, Wright, 1984b):

$$F_y = \frac{H_o - H}{H_o}$$

where H_0 is the conditional probability that *y* is heterozygous and *H* is the unconditional probability that *y* is heterozygous at a specific locus.

4.3 Results and Discussion

4.3.1 Pedigree map size, complexity, and offspring family distribution

The largest number of generations traced was eight and the minimum was 0. Of the 3291 accessions, 381 (11.58%) accessions had full parentage, 2468 (74.99%) accessions had one parent listed, and 442 (13.43%) had no parentage listed. There were 442 (13.43%) orphan lines, 2348 (71.35%) terminal lines, and 3230 relationships were identified in the pedigree. Visual inspection of the pedigree did not suggest any obvious bottlenecks.

Of the 20.44% of accessions that had recorded offspring families, the number of offspring families ranged from 1 to 356. The mean offspring family number across the whole population was 1.845. Within the accessions that had offspring families, the mean number of offspring families was 2.045 and the mode number was 1. In the overall pedigree map, 3.73% of the accessions had only one offspring family, another 3.73% had two offspring families, 1.61% had three offspring families, and 1.21% had four offspring families. We also identified accessions with large numbers of offspring families. The largest percentage of accessions with offspring families was 3.79%, which had between 10 and 30 families.

These four accessions had >80 offspring families: F510, F709, F1139, and F2508. The first three accessions (F510, F709, and F1139) were Grasslands cultivars, whereas accession F2508 was a breeding line known as “multi-leaf selection” and was introduced into the MFGC in 1990. The accession F510 (k = 0.019) had 265 full-sib and half-sib offspring families, and the earliest traceable accession was introduced into the MFGC in 1952. F709 (k = 0.014) had 183 full-sib and half-sib offspring families and the earliest traceable accession was introduced into the MFGC in 1960. F1139 (k = 0.019) had 81 half-sib offspring families, and its earliest traceable accession was introduced into the MFGC in 1968. Accession F2508 (k = 0.057) had 356 full-sib and half-sib offspring families. Because three of the four accessions were associated with commercially available cultivars, the populations must have been phenotypically desirable.

In the 1930s, a significant number (529) of accessions was associated with germplasm collected in New Zealand (Figure 4.2). The New Zealand accessions peaked in the 1940s but declined steadily thereafter until the 1990s and 2000s. Comparatively, the 1970s was a decade where the largest number of accessions (483) was imported from foreign countries to New Zealand. The countries of collection that contributed the most were: the United States (208), the former Soviet Union (65), Germany (34), England (28), Sweden (19), Poland (15), Belgium (13), and France (13).

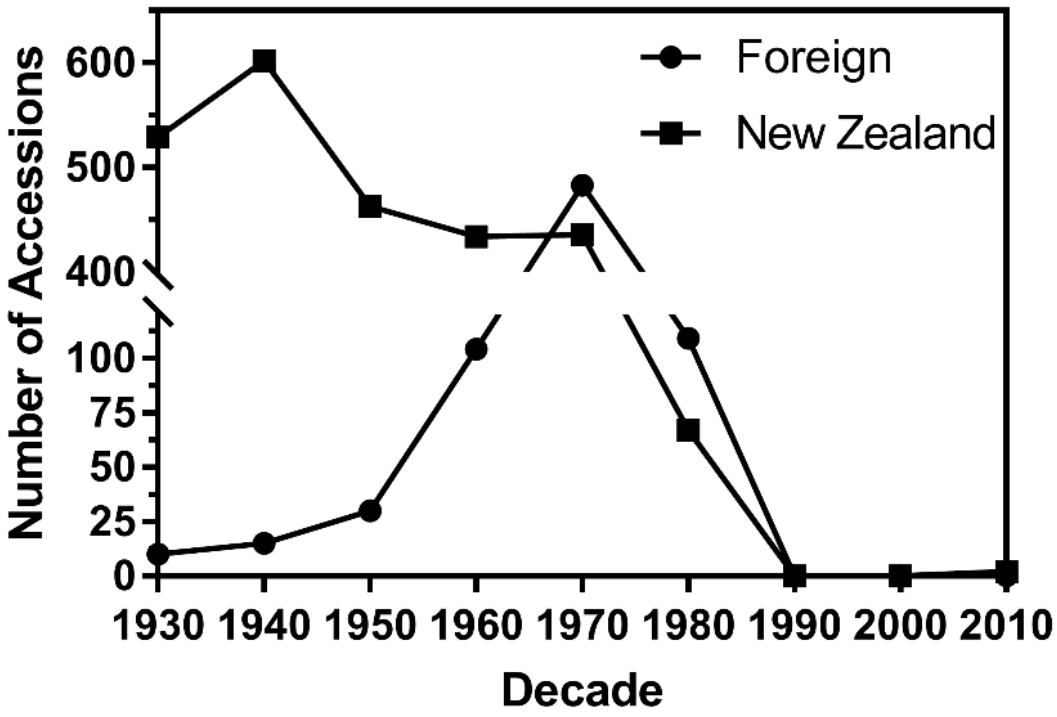

Figure 4.2 The number of *Trifolium pratense* L. accessions associated with foreign countries and the number of accessions collected in New Zealand between 1930 and 2010.

4.3.2 Founding ancestors and important introductions

Founders

A total of 30 founding accessions was identified in the red clover germplasm database. The founders are defined here as the oldest accessions with no parentage recorded, and they were introduced into the MFGC between 1934 and 1937. Of these, 25 were classified as 'New Zealand Broad', two were classified as 'English Broad', two were classified as 'English Giant Hybrid', and one was classified as 'Cotswold Broad'. The first importations of red clover into New Zealand were mainly from English commercial firms.

'Broad red clover', an early flowering cultivar that lacks persistence, was widespread in the drier regions of the South Island and was used as the base population in breeding programs (Corkill, 1949, Wratt, et al., 2015). By 1941, the first synthetic red clover had been formed, and some parent plants were derived from 'English Broad' and 'English Giant Hybrid'. The synthetic red clover was increased, subjected to agronomic trials and labelled as 'New Zealand Broad Red Clover'. This was later renamed in 1964 as 'Grasslands Hamua'.

Influencing ancestors

There were five accessions that strongly influenced the population structure of red clover (Figure 4.3b), namely F2508, F3540, F3801, F510, and F1139. Accession F2508 was listed in July 1990 and

had one accession listed for parentage, F2367. Accession F2367 was collected from Turkey in August 1988. The climate of Turkey is typical of Mediterranean climate, with hot dry summers and mild to cool wet winters. Traditionally, Mediterranean phenotypes have been desirable in New Zealand, as they exhibit drought tolerance traits. Its climatic origins could be one of the reasons why this accession influenced the population structure (Figure 4.3b).

We hypothesize that adaptation to the target environments in New Zealand is a driving force behind the influence of the founding accessions. Here, we describe six relationships. Accession F3801 was listed in January 2012 and had the half parentage listed as F3540 (Figure 4.3b). Accession F3540 was listed in April 2008 and arose out of a biparental cross between F2903 and F2499. Accession F2903 is a Grasslands cultivar, 'Grasslands Sensation' (Claydon, et al., 2003), which was traced back to a polycross from various accessions. 'Grasslands Sensation' is an early-flowering diploid cultivar, which persists under grazing. Accession F2499 can be traced back to an introduction from Portugal (F2414). Two other accessions, one known as 'Renova', was an introduction from Belgium (F2140), and one known as 'Changins' was an introduction from Switzerland (F2138).

Accession F510 (Figure 4.3b) was listed in April 1952 and was a progeny of accession F480. Accession F510 was defined as a highly contributing parent, with 265 offspring families. Accession F480 had no listed parentage. It was brought into MFGC in January 1950 and collected as a naturalized accession in the Manawatu region of New Zealand. Accession F1139 produced 81 offspring families and was listed in January 1968. The half parentage listed was F1002 or 'Grasslands Hamua', a commercial cultivar.

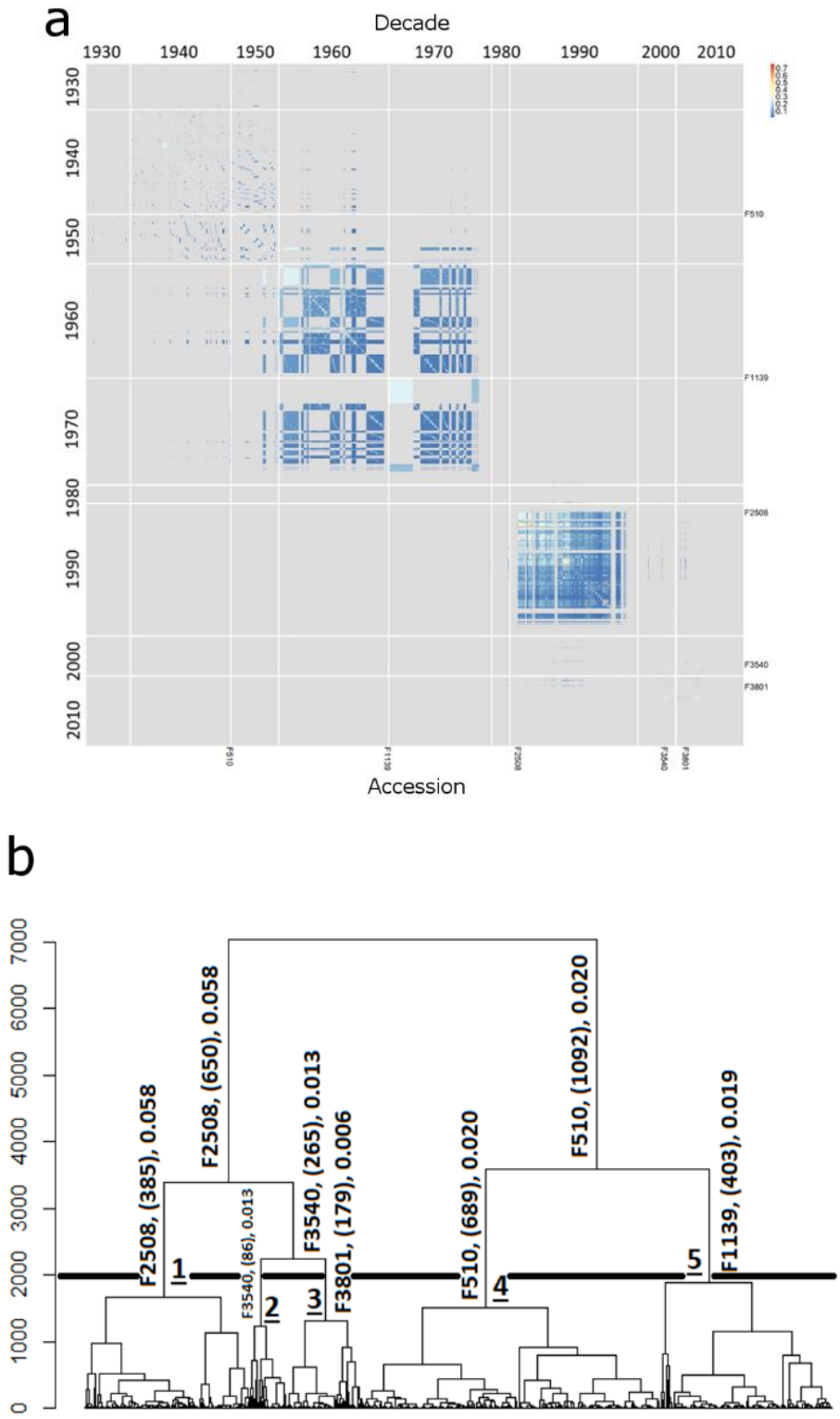

Figure 4.3 (a) Kinship heat map of *Trifolium pratense* L. accessions at the Margot Forde Germplasm Centre. (b) Dendrogram of accessions included in the study. Accessions on the branches of the dendrogram at the fusion points are influencing ancestors of each cluster with the highest mean kinship in the cluster. The number of accessions in each cluster are indicated in parentheses. The solid line indicates the distance used to separate the different clusters studied. Clusters are numbered and underlined for reference.

4.3.3 Diversity and inbreeding

Diversity

The overall mean kinship from 2,201,851 pairwise combinations within the species, with kinship values ranging from 0 (no relatedness) to 1 (full relatedness), was $k = 0.005$. A total of 498 accessions (28.59%) had a mean kinship of 0, whereas 1244 accessions (71.41%) had kinship >0, with a maximum value of $k = 0.0576$. These kinship values indicate that there are low levels of relatedness among accessions in the red clover collection (Figure 4.3a).

The cumulative trend of kinship across the decades shows an increase in kinship of 0.021 from 1930 to 1970, until a plateau occurred between 1970 and 1980. A sharp increase of 0.012 occurred between 1980 and 1990. Between 1990 and 2010, there was a small increase of 0.001 (Figure 4.4).

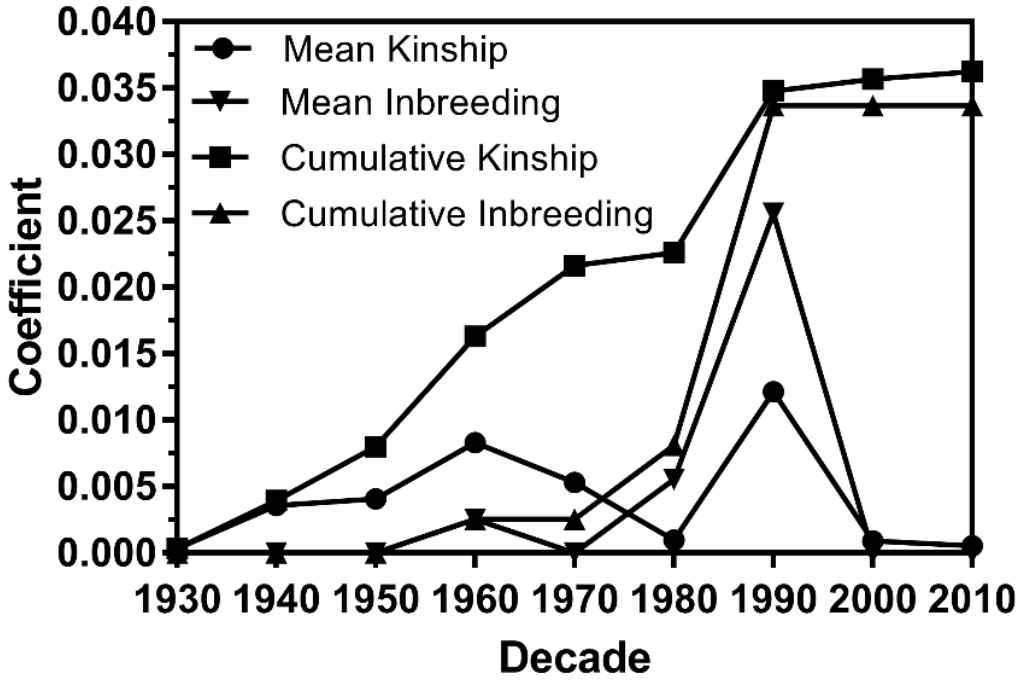

Figure 4.4 The trend in cumulative and mean kinship and inbreeding in *Trifolium pratense* L. accessions at the Margot Forde Germplasm Centre between 1930 and 2010.

The dendrogram (Figure 4.3b) has five clusters in the population. The clustering and identification of current commercial cultivars in the clusters can provide insight into the diversity present in the market and is of particular interest to the industry.

Parts of the dendrogram can be explained by groups of accessions with common countries of origin. Clusters 2 (86 accessions), 3 (179 accessions), 4 (689 accessions), and 5 (403 accessions) had some, not all, accessions associated with country of origin. Cluster 2 had 15 accessions from Switzerland. Cluster 3 had five countries associated with accessions: Switzerland (13), Turkey (6), Spain (20), Portugal (4), and Argentina (1). Cluster 4 had 30 accessions from Turkey. Cluster 5 had seven

countries associated with accessions: Portugal (4), Spain (10), Armenia (22), Georgia (30), Tajikistan (18), Greece (24), and Russia (13).

Clusters 4 and 5 were the two clusters from which commercial cultivar data could be extracted. Cluster 4 had 16 (2.32%) accessions associated with the "Grasslands cultivar" group. Accessions in Cluster 4 were introduced between October 1937 and January 1965. The geographic origin of commercial accessions within Cluster 4 was confined to New Zealand. However, 15 accessions were linked to the Manawatu region (North Island) and one to the Canterbury region (South Island). Cluster 5 had five accessions associated with "Grasslands cultivars" and three accessions were associated with "other cultivars." The accessions in Cluster 5 were introduced between January 1968 and July 1982, notably later than those in Cluster 4. 'Grasslands Hamua', 'Grasslands Turoa', and 'Redwest' (Charlton and Stewart, 1999) are the three cultivars present in Cluster 5. 'Grasslands Hamua' and 'Grasslands Turoa' have similar soil and moisture requirements. However, 'Grasslands Turoa' is more persistent. 'Redwest' is a cultivar that was reselected from 'Grasslands Hamua' on the basis of low estrogen content. Its general agronomic traits are very similar to 'Grasslands Hamua'.

An interesting observation is the high number of Turkish introductions in Cluster 4. In contrast, a variety of countries contributed towards the composition of Cluster 5, with fewer accessions recorded as domestic cultivars than Cluster 4.

Separation between Clusters 4 and 5 might be attributable to selection pressure. There were five records of the cultivar 'Grasslands Hamua' used as a parent. These accessions were split between Clusters 4 and 5, with one accession in Cluster 4 and four accessions in Cluster 5. The accession in Cluster 4 originated from the Canterbury region, whereas all but one of the accessions in Cluster 5 originated from Manawatu. Accession F1139 (Figure 4.3b), the most influencing ancestor of Cluster 5, is a progeny of 'Grasslands Hamua'. The separation among clusters and the presence of released cultivars within different clusters suggests that breeding and selection for different target environments in New Zealand might have played a role in the genetic divergence observed.

4.3.4 Inbreeding, indirect relationships, and unrelated accessions

The mean inbreeding level was 0.56%; among the accessions with inbreeding coefficients >0, the mean was 10.68%. The frequency of the levels of inbreeding within the inbred accessions peaked at coefficients 0.07 and 0.13. Among the accessions with inbreeding coefficients >0, the lowest level of inbreeding was 0.015 and the highest level was 0.25. Across the whole dataset, 1651 accessions (94.78%) had an inbreeding coefficient of 0, whereas 91 accessions (5.22%) showed some inbreeding.

The combined trends of inbreeding and kinship in the red clover collection are shown in Figure 4.4. Although the inbreeding levels are low, they seem to have increased across time. A correlation was found among average inbreeding per decade and the number of introductions per decade ($r = 0.62$). The trend in inbreeding from the decades of 1930 to 2010 shows no inbreeding from the 1930s to the

1970s, except for a slight increase (0.002) in the 1960s. The highest inbreeding occurred between 1980 and 1990, where it increased to 0.025. There was no further increase in inbreeding in subsequent decades (Figure 4.4). The overall inbreeding value was 3.37%.

Although red clover is difficult to self, it is easily inbred (Taylor, et al., 1980). Sib-crossing can result in severe inbreeding depression and populations, with inbreeding via selfing cannot continue beyond two to three generations. If selfing reaches past these generations, loss of vigor, fertility, and viability of pollen occurs. Because of the sensitivity of red clover to inbreeding, it is important to breed it through populations, rather than individual plants (Taylor, et al., 1996a).

Although inbreeding is regarded as deleterious, it should be noted that if crosses are performed correctly, inbreeding can lead to increased genetic gain. Taylor, et al. (1970) investigated the effects of one generation of selfing on pseudo-self-compatibility (PSC) and seed and forage traits to assess the degree of hybrid vigor and its correlation to persistence. It was concluded that inbred parental lines might be maintained through either vegetative reproduction or by seed through PSC. Although selfing reduced the yield and persistence rates, these were regained in single crosses. This was strengthened by Duncan, et al. (1973), who found that PSC decreased with inbreeding. They proposed a new method of inbred line maintenance, combining the favorable features of both vegetative and seed maintenance of the inbred lines. This ultimately resulted in single cross seed from different clonal sources for the production of double-cross hybrid red clover.

4.3.5 Germplasm centers and the effect of forage breeding strategies

The importance of legumes is recognized by the large number of germplasm collections worldwide, with >1,000,000 accessions. Approximately 74,100 accessions of *Trifolium* are held in global collections; 53% are wild, and 14% are cultivated (Smýkal, et al., 2015).

Kouamé, et al. (1993) evaluated >800 accessions of red clover, originating from 41 countries and held in the Germplasm Resource and Information Network of the NPGS, USDA. Their analysis found a large range of diversity across all countries, with the highest diversity contained in accessions originating from northern and eastern Europe. That study showed that clustering resulted from similar agronomic traits and it provided information for the creation of a core collection and more effective utilization of red clover germplasm. This is comparable with our study, as we also found a large range of diversity.

Mosjidis, et al. (2006) assessed the biochemical diversity present in red clover accessions of the USDA NPGS core subset, concluding that genetic diversity was high and that there was nearly double the variability in wild populations compared with cultivars or landraces. Dias, et al. (2008) evaluated the diversity of the same core collection of 85 accessions of red clover, at both the morphological and molecular levels. An analysis of molecular variance showed that 83.6% of the variation was contained

within population. This is valuable for breeding programs to use within and between population variation to breed for improved varieties.

The rate of genetic gain in red clover is low because of the long breeding cycles, complex outbreeding mating systems that suffer inbreeding depression, and interaction of genotypes with environmental factors (Annicchiarico, et al., 2015). Although there has been a significant rise in the number of genetic studies in forages worldwide to suggest novel breeding strategies, the basis of forage plant breeding still largely relies on recurrent phenotypic selection. Mass selection, recurrent phenotypic selection, polycrossing, and backcrossing are the most common selection and crossing techniques in red clover worldwide (Riday, 2010, Taylor, 2008).

Breeding strategies for red clover are based mainly on mass selection, suggesting that it should be the first technique deployed in a breeding program to maintain a large population and diversity (Taylor, et al., 1996a). Mass selection has been used largely for pest resistance and persistence in red clover (Riday, 2010). Polycrossing has been sparsely used in red clover; however, progeny testing and multiple locations must be used to account for genotype × environment interactions (Taylor, et al., 1996a). In the red clover pedigree dataset, 183 accessions were identified as bred by polycross and 3108 were bred by biparental crosses. Polycrosses were identified for the decades of 1950, 1960, and 1970, with 47, 33, and 103 polycrosses, respectively.

Backcrossing is seldom used in red clover breeding, mainly because of the promotion of inbreeding depression. However, when it is used, it is in breeding programs to incorporate disease or pest resistance. Taylor, et al. (1986) used backcrossing to incorporate resistance genes to a strain of Bean yellow mosaic virus in the cultivar ‘Kenstar’. However, to date, there are no other cultivars of red clover that have been developed via the backcross method.

Unlike white clover, using related species in red clover to increase the genetic diversity has had limited success. Cleveland (1985) summarized the efforts made in hybridizing red clover with related species. However, because of most of the relatives are annuals, the perennial trait proved hard to maintain. Sterility in the progeny has been one of the biggest problems in this regard, and it is the prime reason why no commercial cultivars are available on the market for *Trifolium* ISHs. Forage species retain many of the traits of their wild progenitors because their domestication history (<3000 years) is more recent than other crops (>10,000 years) (Walton, 1971) and, in general, forage species are considered to be only partially domesticated (Gepts, 2004). This is shown particularly with persistence and dry matter yield under grazing, especially in New Zealand, and is one of the primary reasons why progress in breeding for complex fitness traits has been slow.

Nutman, et al. (1981) evaluated the symbiotic effectiveness of four cultivars of red clover and found that crosses between cultivars were more effective than those within cultivars, indicating heterotic

effects. This shows the opportunity for breeding across gene pools, rather than within the same gene pool.

The current breeding strategies relative to the red clover accessions in the MFGC have been successful thus far, as inferred from the metrics reported here, such as number of accessions bred and generated, inbreeding and kinship levels, and the number of cultivars relevant to specific sectors of genetic diversity. The potential for breeding across newly identified breeding pools and continuing research into germplasm diversity will provide new techniques to widen the breeding pool.

4.4 Conclusions

Red clover genetic resources held at the MFGC have been valuable for germplasm diversity, and this is reflected in low kinship and inbreeding levels. International collection expeditions and germplasm exchanges have proven successful in achieving diversification and introduction of new traits into cultivars in New Zealand and overseas. Influencing ancestors have resulted from these introductions. The newly identified population patterns and divergent clusters of germplasm will allow more informed and targeted breeding decisions. Improved populations with new additive variation can be obtained via crossing among accessions with disparate relatedness; these are considered untapped resources that will increase germplasm utilization for pre-breeding and cultivar development.

Chapter 5

Pedigree analysis of pre-breeding efforts in *Trifolium* spp. germplasm in New Zealand

This chapter has been published in BMC Genetics:

Egan, L.M., R.W. Hofmann, P. Seguin, K. Ghamkhar and V. Hoyos-Villegas. 2020. Pedigree analysis of pre-breeding efforts in *Trifolium* spp. germplasm in New Zealand. BMC Genetics. doi: 10.1186/s12863-020-00912-9

5.1 Introduction

The earliest recorded use of legumes is in the grasslands of the Mediterranean basin and today they are used in agricultural pasture systems (Knowles, et al., 2003). The importance of genus *Trifolium* was recognised very early on by naturalists and herbalists. It was previously defined as a much broader genus and included two other genera, *Lotus* and *Melilotus* (Ellison, et al., 2006).

To date, the *Trifolium* genus of the *Fabaceae* family is made up of 250 species (Ansari, et al., 2004, Frame, et al., 1986, Gillett, et al., 2001, Lewis, 2005, Maxted, et al., 2001, Scoppola, et al., 2018, Zohary, et al., 1984). They are distributed throughout the temperate and subtropical regions of both northern and southern hemisphere, particularly in Europe, northwest and central Asia, northeast Africa, parts of sub-tropical Africa and South Africa, western North and South America, Australia and New Zealand. Approximately 25 species are of significance as feed for ruminant animals, of which sixteen are economically important (Frame, et al., 1986, Gillett, et al., 2001, Speer, et al., 1985). Russel, et al. (1976) state that the *Trifolium* species are among the most important and valuable forage legumes in the world. Nitrogen fixation is a quality which has driven the use of *Trifolium* species in pastoral systems. However, many *Trifolium* species are underutilised in agricultural systems or their use is yet to be defined in agriculture (Maxted, et al., 2001).

The Margot Forde Germplasm Centre (MGFC) is New Zealand's national germplasm centre for grasslands plant species. The role of this genebank is to collect, replenish, conserve and distribute accessions of forage species to be used for research or breeding. The collection contains over 65,000 wild accessions from more than 100 countries, comprising 2200 species from 350 genera and over 70 plant families from wild collections, foreign and domestic cultivars, breeding lines and genetic stocks. There are six *Trifolium* species that are well represented with pedigree information in the MFGC database. These are *T. arvense, T. subterraneum, T. ambiguum, T. dubium, T. hybridum* and *T. medium*. These species have been subject to breeding programmes from the 1950s, and some species have been hybridised to breed *Trifolium* ISHs. The ISH programmes have used white clover as the recurrent parent and hybridised it with a closely related species to improve targeted traits in white

clover. For example, *T. occidentale* is well-adapted to dry habitats and could be a potential source of drought-tolerant genes for white clover (Hussain, et al., 2013).

The *T. occidentale* x *T. repens* ISH is an important breeding programme for the improvement of white clover in New Zealand. Another species, *T. arvense*, commonly known as rabbitfoot clover, is an annual clover that is native to most of Europe, excluding the Arctic zone, and western Asia. It can grow in a broad range of soil types, but prefers sandy or non-irrigated land (Pritchard, et al., 1988). It is used in short-lived pastoral systems and low fertility pastures, often in dry hill country environments (White, et al., 1999). Hancock, et al. (2012) used GM to integrate the transcription factor, *TaMYB14*, from *T. arvense* into *T. repens*. The GM product has the potential to decrease methane emissions and reduce bloating in livestock. This is due to the increased level of proanthocyanidins which reduce the level of protein degradation in the rumen, decreasing gas and foam formation (Aerts, et al., 1999).

Another clover species, *T. subterraneum*, or subterranean clover, is an annual species native to the Mediterranean region, West Asia and the Atlantic coast of Western Europe. Subterranean clover is sown in over 29 million hectares worldwide and has the most contribution to livestock feed production among all annual clovers (Kaur, et al., 2017). Subterranean clover has a unique characteristic where it buries its seeds so that the seed development occurs underground. This specialty enables subterranean clover to thrive in poor quality soil and drought regions, making it a viable option for dryland farmers (Bennetts, et al., 1946, Nichols, et al., 2013a). Subterranean clover is one of the most commonly grown forage crops in Australia due to its ability to withstand the extreme drought and soil types, and is a source of high quality forage (Bennetts, et al., 1946, Nichols, et al., 2013a). Since the introduction of subterranean clover into Australia, more than 40 cultivars have been bred and released (Nichols, et al., 2013a). Subterranean clover is a diploid (2n = 2x = 16), mostly self-pollinating species, with a genome size of 540Mpb (Abberton, et al., 2005, Dudchenko, et al., 2018, Ghamkhar, et al., 2012, Hirakawa, et al., 2016, Kaur, et al., 2017).

A genetically diverse species, *T. ambiguum*, most commonly known as ‘Caucasian’ or ‘Kura’ clover, is a species native to Asia and the Caucasus (Armenia, Ukraine, Turkey and Iran) and was introduced to North America and Australasia (Bryant, 1974, Taylor, et al., 1997). It is a rhizomatous perennial, found naturally up to high altitudes and is adapted to a wide range of environmental conditions (Dear, et al., 1985, Williams, et al., 2011). The large root and rhizome mass and persistency of Caucasian clover has made it desirable in agricultural environments exposed to extreme heat, drought and cold. However, the slow establishment rate, very specific rhizobial requirements, and inability to produce commercially viable amounts of seed has decreased the appeal of sowing it in a pasture system (Bryant, 1974, Maxted, et al., 2001). Trials in the high country of the South Island of New Zealand have shown that, in comparison to white clover, *T. ambiguum* increases the legume content of a pasture in competition with grasses, and therefore potential nitrogen fixation in paddocks (Charlton, et al., 2006). *Trifolium ambiguum* exists in diploid (2n = 16), tetraploid (2n = 32) and hexaploid forms

(2n = 48). The ploidy of the species affects traits such as flowering date and persistence but is not directly related to overall yield (Bryant, 1974, Dear, et al., 1985). Diploids are often the first to flower and are more persistent than tetraploids and hexaploids (Dear, et al., 1985, Widdup, et al., 1996).

A group of three other species, namely, *T. dubium, T. hybridum and T. medium* are less common in pastoral systems and have been predominantly used in research. The first species, also known as 'Suckling clover', is native to Europe and is a cross between *T. campestre* and *T. micranthum* (Hedlund, et al., 2003). It is an allotetraploid species (2n = 4x = 30) (Bulińska-Radomska, 2000). Alsike clover or *T. hybridum* is a clover that originates from continental Europe but has established in the British Isles and throughout the temperate regions of the world. It is often grown for hay or silage, is highly self-sterile and unlike the name suggests, is not of a hybrid origin (Williams, 1951). In New Zealand, it is used often in the South Island hill country for pasture and hay. It is adaptable to a wide range of conditions and has rapid establishment (Widdup, et al., 1994). The last species in this group, *T. medium*, commonly known as 'Zig-zag clover', is a native European species and is similar in appearance to red clover but with narrower leaflets and no white leaf markings (Choo, 1988). It is a rhizomatous perennial clover with long persistence and has the ploidy of 2n=10x=80 (Isobe, et al., 2002, Merker, 1984).

Although natural interspecific hybridization is uncommon in *Trifolium*, there have been several studies showing that it is possible (Brewbaker, et al., 1953, Malaviya, et al., 2004, Marshall, et al., 1995, Williams, et al., 2011). The ISH breeding programmes within the genus commenced over 50 years ago. Two common objectives from these breeding programmes were to understand the evolutionary relationships within the genus and to introgress desirable traits into the species (Abberton, 2007). *Trifolium occidentale* is a diploid (2n = 16), perennial stoloniferous species that is closely related to white clover (*Trifolium repens)* (Abberton, 2007, Williams, et al., 2011). It is indigenous to the coastal areas of Portugal, Spain, France and the British Isles, hence its tolerance to saline and dry habitats. This trait provides a potential source of drought-tolerant genes that could be used to improve white clover (Hussain, et al., 2013). The ISH of *Trifolium* species crossed with white clover will allow elite germplasm to be bred with alleles that are not present in white clover populations (Abberton, 2007). Pederson, et al. (1989) performed a variety of crosses between *Trifolium* species, and *T. occidentale* x *T. repens* hybrids were the only fertile hybrids, which also showed resistance to peanut stunt virus. However, often the two species do not cross easily and result in near-sterile triploid hybrids (Hussain, et al., 2013). Using 4x *T. occidentale* has yielded more successful crosses, resulting in significant advances in the introgression of drought and salt tolerance traits to white clover (Chou, et al., 1968, Gibson, et al., 1969, Pederson, et al., 1989, Williams, 2014).

Pedigrees are used in plant breeding to visualise the breeding crosses and the transmission of alleles responsible for trait expression and consequent breeding patterns (Shaw, et al., 2014). They are a crucial first step in identifying genetic bottlenecks in breeding populations, and integration with

genomics has increased their relevance in plant breeding programmes even further (Acquaah, 2012). The use of plant pedigrees is proving valuable for to lifting the rate of genetic gain in white clover (Hoyos-Villegas, et al., 2019), increasing the range of environmental adaptation and increasing tolerance to plant stressors. Pedigrees are also used alongside molecular studies to increase the accuracy of molecular phylogenetic studies (Crossa, et al., 2010). Pedigree maps can also be used to identify germplasm variation which can be utilised in future pre-breeding decisions (Egan, et al., 2019a, Egan, et al., 2019b). For this to occur effectively, knowledge of population structure and relatedness coefficients are the first steps (Keller, et al., 2002).

Pre-breeding is becoming an increasingly important prerequisite of plant breeding programmes. Plant species that focus on new variation and flow of allelic variation benefit from pre-breeding research. Related species that with desirable traits in the genus *Trifolium* are often used in pre-breeding efforts and have been increasingly utilised in recent years. We used historical pedigree data from the *T. ambiguum, T. arvense, T. dubium, T. hybridum, T. medium, T. subterraneum* and the *T. repens* x *T. occidentale* ISH collections held at the MFGC in Palmerston North, New Zealand. The objectives of this study were (i) to create a pedigree map for each *Trifolium* species (hereafter referred to as *Trifolium spp.*), (ii) to analyse patterns affecting inbreeding and kinship, and (ii) to investigate variation in the collection that can be potentially utilised in future pre-breeding work.

5.2 Materials and Methods

5.2.1 Germplasm

The *Trifolium* spp. that were selected for this study were selected based on historical breeding activity and their importance in New Zealand's pastoral systems. The formal identification of the germplasm used in this study was undertaken by the MFGC. The germplasm is available in the MFGC. A total of six species: *T. arvense, T. subterraneum, T. ambiguum, T. dubium, T. hybridum* and *T. medium* have been the subject of active breeding programmes since the 1950's. The *T. repens* x *T. occidentale* ISH breeding programme has pedigree information available from 2015. International collection trips to collect germplasm from abroad in these species have been successful in bringing new germplasm back to the MFGC. Where necessary, permits were obtained for collections. Whilst these species are currently not major pastoral species used in New Zealand, some are often used in pastoral mixes or have been critical in research for improving the major species *T. repens*. One species, *T. ambiguum*, and one hybrid, *T. repens* x *T. occidentale*, have been actively used in the *Trifolium* ISH breeding programme and have improved root systems in *T. repens*. These species were chosen as (i) they are often used in pastoral systems as minor pastoral species or, (ii) they are used thoroughly in pastoral research as a research species or, (iii) they can be hybridised with a major pastoral species.

5.2.2 Data filtering

Pedigree map construction was undertaken using pedigree data from the MFGC database. Seven minor *Trifolium* spp. were used in the pedigree analysis, chosen based on the size and completeness of the data available (Table 5.1). The methodology used in this study is the same as in Egan, et al. (2019a) and Egan, et al. (2019b). In short, accessions have been introduced over a range of decades, from 1950 to 2010. A range of breeding techniques have been used during population development over this period, including bi-parental cross and polycross methods. The pedigree maps were constructed using Helium, a software that allows the visualisation of large pedigrees (Shaw, et al., 2014). Accessions with large numbers of offspring families were identified. These largely contributing accessions were determined as parents with several progeny well-above the mean number of progenies for each species. The term "accession" here is used to refer to any seed material entered in the MFGC as a distinct population with an identification number.

Table 5.1 Completeness of parentage information of germplasm of seven *Trifolium* species from the Margot Forde Germplasm Centre, New Zealand. Half parentage indicates that one parent is listed.

Species	Number of accessions in pedigree map	Number of accessions used in parameter analysis (% from total)	Full parentage (% from total)	Half parentage (% from total)	Null parentage (% from total)
T. ambiguum	814	772 (94.8%)	0 (0.00%)	772 (94.8%)	42 (5.2%)
T. arvense	88	55 (62.5%)	0 (0.00%)	56 (63.6%)	32 (36.4%)
T. dubium	214	191 (89.3%)	1 (0.5%)	190 (88.8%)	23 (10.8%)
T. hybridum	364	196 (76.6%)	1 (0.3%)	297 (81.6%)	66 (18.1%)
T. medium	112	107 (95.5%)	29 (25.9%)	78 (69.6%)	5 (4.5%)
T. subterraneum	473	248 (52.4%)	0 (0.00%)	248 (52.4%)	225 (47.6%)
T. repens x *T. occidentale*	1472	643 (43.7%)	579 (39.3%)	594 (40.4%)	299 (20.3%)

A smaller number of accessions were used in the derivation of population parameters than the number used to construct the pedigree map, because of the completeness of pedigree information available. The accessions were categorised based on identifiable parental information, full or half parentage indicated, and the type of cross that was conducted (e.g., biparental crosses). Polycrosses represented a small number of accessions and were excluded from the parameter subset. This was due to the software package not having an option for handling the pedigree data with both biparental crosses and polycrosses. Further, polycrosses do not fit the allele frequency expectations of biparental crosses and could therefore skew estimates of relatedness as indicated in Egan, et al. (2019a) and Egan, et al. (2019b).

5.2.3 Data analysis

The R package 'pedigree' was used to calculate the number of offspring, kinship and inbreeding (Coster, 2015).

The pedigree information was used to calculate kinship (*k*) by using a recursive application of two formulae:

$$[1] \qquad F_{yy} = \frac{(1+Fm_y f_y)}{2}$$

$$[2] \qquad F_{xy} = \frac{(F_{xmy}+F_{xfy})}{2}$$

where *F* is the coefficient of inbreeding, F_{yy} is the kinship of *y* with *y* and F_{xy} is the kinship of two individuals, assuming *x* is not a progeny of *y*. When *x* and *y* are founding accessions, F_{xy} =0. The random inheritance of one gene from each parent is transcribed as *my* and *fy* of *y*, and the relationship of these genes is $Fm_y f_y$ (Fernando, et al., 2006). *k* is the level of relatedness of the accession.

A dendrogram was used to visualise and summarise the clustering of the populations based on pedigree information and kinship coefficients. The influencing ancestors were termed by identifying the accessions with the highest mean kinship per cluster. This was verified by pedigree lineages.

Inbreeding was calculated using the formulae (Crow, et al., 1970, Wiggans, et al., 1995, Wright, 1984a):

$$F_y = \frac{H_o - H}{H_o}$$

where *H* is the unconditional probability that *y* is heterozygous at any given locus. H_0 is the conditional probability that *y* is heterozygous at a given locus, when the genes are not identical-by-descent (Fernando, et al., 2006).

5.3 Results

5.3.1 Pedigree map sizes and complexity

Between 0 and 3 generations were traced among *Trifolium* spp., with *T. subterraneum* and *T. repens* x *T. occidentale* having most generations. The least number of generations (0-1) was observed in *T. arvense* and *T. medium* (Table 5.2). The shallowness of the pedigree maps is due to the low number of generations and reflects low breeding activity. The completeness of parentage was variable (Table 5.2) with *T. repens* x *T. occidentale* having the greatest number of accessions with complete parentage (39.3%). *T. subterraneum* had the greatest number of accessions with null parentage (47.6%). Terminal lines are described as accessions that are not involved in any further breeding and reflect on the amount of breeding activity. with the highest number of terminal lines in the pedigree (765

accessions, 94% of total) belonging to *T. ambiguum*, and the smallest number ((235 accessions, 50% of total), being in *T. subterraneum*. The breeding activity peaked in the 1990's due to an influx of accessions being deposited at the MFGC. This is from both breeding crosses and accessions introduced from international collection trips. Visual inspection of the pedigree maps did not suggest any bottlenecks (data not shown).

The smallest range of offspring distribution (0-3) was in *T. subterraneum*, compared to the largest distribution of *T. ambiguum* (0-122) (Table 5.2). Within accessions that had offspring, the highest average number of offspring (39) belonged to *T. ambiguum*, indicating a high level of breeding activity. This contrasts to *T. subterraneum*, *T. hybridum* and *T. dubium* which had an average offspring number of 1, showing low breeding activity.

Accessions with large numbers of offspring families were also identified. Only one accession of *T. ambiguum*, 'AZ2640' (*k*=0.025), contributed 34 half-sib and full-sib offspring families to the population. This accession had the half parentage of 'AZ1981', which was an introduction from New South Wales, Australia in 1988 and is known as cultivar 'Monaro'.

A total of four accessions of *T. dubium*, AZ1840, AZ170, AZ2022, and AZ1649, contributed 93 progeny (44%) to the population. Accession AZ1840 was an introduction into the MFGC in 1984 from the Manawatu region of New Zealand and produced 20 half-sib and full-sib offspring families. Accessions AZ170, AZ2022 and AZ1649 contributed 22, 22 and 29 half-sib and full-sib offspring families respectively. Accession AZ2022 was an introduction from the Manawatu region, and accession AZ1649 was an introduction from Portugal in 1983.

Two accessions, AZH1605 and C25897, contributed a large number of half-sib and full-sib offspring families to the *T. repens* x *T. occidentale* ISH, population. Accession AZH1605 (k=0.008) was known as a $BC1F_2$ hybrid with accession AZH784 as its parent. Accession C25897 was bred in 2010 with accession C25638 as its parent. The white clover cultivars, 'Grasslands® Mainstay' and 'Grasslands® Kopu II', were listed among the lineage. Mainstay white clover is a large-leaved, high yielding white clover that has high dry matter yield (Agricom, 2015). Kopu II is a large-leaved, high yielding white clover which has a high tolerance to clover root weevil (Woodfield et al., 2001). as the aim of the ISH was to increase yield, tolerance to pests and drought tolerance, it is sensible that these two cultivars were among the parents.

Table 5.2 Number of generations, offspring distribution and number of terminal and orphan lines for seven *Trifolium* species at the Margot Forde Germplasm Centre, New Zealand.

Species	Number of generations	Offspring distribution	Number of accessions with offspring	Non zero average offspring range	Non zero average offspring	Number of orphan lines	Number of terminal lines	Year of entry of the first accession
T. ambiguum	0-2	0-122	6	2-122	39	42 (5.2%)	765 (94.0%)	1962
T. arvense	0-1	0-5	4	1-5	2	32 (36.4%)	52 (59.1%)	1962
T. dubium	0-2	0-25	31	1-25	1	24 (11.2%)	160 (74.8%)	1956
T. hybridum	0-2	0-7	27	1-7	1	66 (18.1%)	270 (74.2)	1955
T. medium	0-1	0-39	14	1-13	5	5 (4.5%)	93 (83.0%)	1939
T. subterraneum	0-3	0-3	13	1-2	1	225 (47.6%)	235 (49.7%)	1956
T. repens x *T. occidentale*	0-3	0-65	75	1-32	6	732 (49.7%)	1054 (71.6%)	2015

5.3.2 Influencing accessions and introductions

Relevant parents are described in this study as accessions that have structured large portions of the pedigree.

T. ambiguum

Two relevant parents, AZ1981 and AZ104, from *T. ambiguum* influenced the population structure of the species. Accession AZ1981 was introduced in 1985 from New South Wales, Australia. It is better known as cultivar Monaro. Accession AZ104 was also an introduction from Australia in 1962 from the organisation Commonwealth Scientific and Industrial Research Organisation. The introductions from *T. ambiguum* were from Armenia, Georgia, Russia and Turkey (Figure 5.1a).

T. arvense

A total of three relevant parents, AZ1353, AZ2855 and AZ2925 were discovered for *T. arvense*. Accession AZ1353 was introduced into the MFGC in 1979. Accession AZ2855 was introduced in 1989 from the Yugoslavian Forage Legume Collection. Accession AZ2925 was an introduction from Hawkes Bay in 1990 and was collected from verges and cliffs. Hawkes Bay is a region in the east coast of New Zealand's North Island that generally has a warm and relatively dry climate. The northern and central bays of the region include hilly coastal land. The species is often present in areas with sandy soil, sand dunes, and in areas that are not irrigated. The influence of the founding accession AZ2925 could be attributed to the strong adaptation of the accession to an environment similar to its native range. Country of origin was recorded for 47 introductions for *T. arvense* (Figure 5.1b).

T. dubium

A total of two relevant parents, AZ753 and AZ2022, from *T. dubium*. Accession AZ753 was an introduction from Portugal in 1975. Accession AZ2022 was introduced into the MFGC in 1986 and was from a pasture selection from the Manawatu region in New Zealand. The summer climate of Manawatu is temperate, as is the semi-arid climate of Portugal. *T. dubium* had 16 introductions with a recorded country of origin (Figure 5.1c).

T. hybridum

A total of two relevant parents, AB75 and AB230 were identified for *T. hybridum*. Accession AB75 was a Mackenzie country, New Zealand collection and was introduced into the MFGC in 1973. Accession AB230 was collected from Belarus in 1975. Both Belarus and the Mackenzie country have very distinct seasons; long dry summers and cold snowy winters. *T. hybridum* had 20 introductions with a recorded country of origin (Figure 5.1d).

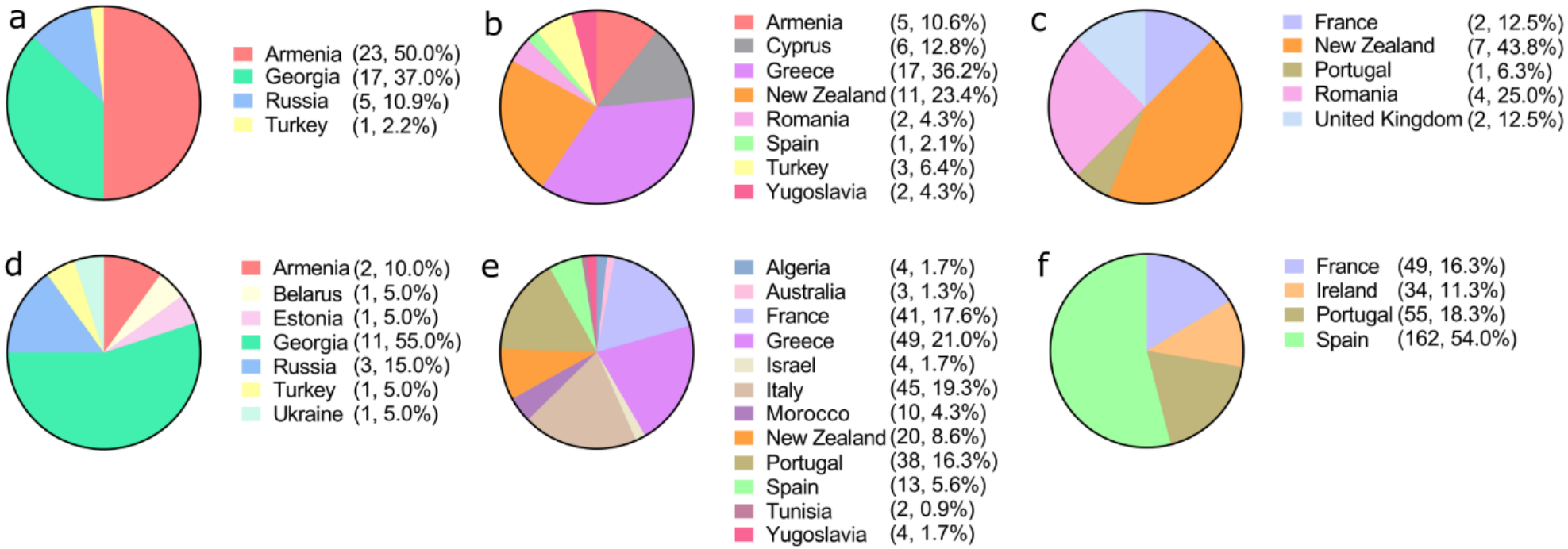

Figure 5.1 Total number of introductions into the Margot Forde Germplasm Centre germplasm collection from listed geographic locations for *T. ambiguum* (a), *T. arvense* (b), *T. dubium* (c), *T. hybridum* (d), *T. subterraneum* (e) and *T. repens* x *T. occidentale* interspecific hybrids (f).

T. medium

A total of two relevant parents, Z1 and Z127, were identified for *T. medium*. Accession Z1 was an introduction from the former USSR in 1939. Z127 was introduced into the MFGC in 1985 and was categorised as a breeding line.

T. subterraneum

A total of four relevant parents, AK1213, AK452, AK799 and AK334, were identified for *T. subterraneum*. Accession AK334 was introduced into the MFGC in 1956 and was an early flowering type. Accession AK452 was introduced into the MFGC in 1962 and was a collection from Morocco. Accession AK799 was a collection from France in 1987 and was subject to a flowering/formononetin breeding selection. Accession AK1213 was collected from South Australia in 1993.

T. subterraneum had the largest number and accessions from different countries. The prominent countries of origin from the highest to lowest number of accessions were Greece (49), France (41), Portugal (38) and New Zealand (20), Spain (13) and Morocco (10) (Figure 5.1e).

T. repens x T. occidentale ISH

The ISH, *T. repens* x *T. occidentale*, had three relevant parents, AZH776, AZH784 and AZH761. Accession AZH776 was introduced into the MFGC and had the commercial cultivar 'Durana' in the parentage (Bouton et al., 2005). Accessions AZH784 and AZH761 entered the MFGC in 2016. The parentage for both accessions are two commercial cultivars, 'Kopu II' and 'Durana'. Kopu II is a large-leaved white clover cultivar that is high yielding and recovers rapidly after grazing (Widdup et al., 2015). Durana is a small-leaved white clover cultivar that originated from the United States (Bouton et al., 2005). A total of 300 parental accessions had country of origin information. The largest number of introductions were made in *T. occidentale* x *T. repens* with 162 parental accessions from Spain and 55 from Portugal (Figure 5.1f).

5.3.3 Diversity and inbreeding

Trifolium ambiguum

There were four accessions that strongly influenced the population structure of *T. ambiguum* (Figure 5.2a); AZ212, AZ2594, AZ3326 and AZ3116. Accession AZ212 (k=0.008) was listed in 1963 and had the half parentage AZ104. Accession AZ2594 (k=0.125) was listed in 1987 and had the half parentage of AZ2359. Accession AZ3326 (k=0.09) was listed in 1996 and had the parentage of AZ212. Accession AZ3116 (k=0.081) was listed in 1993 and had the half parentage of AZ2640. It was known as the 'Monaro late flowering selection'. Monaro is a high performing Australian cultivar, which has however, poor seed yield. However, breeding has progressed using hexaploid Monaro lines to improve the seed yield (Widdup, et al., 1996).

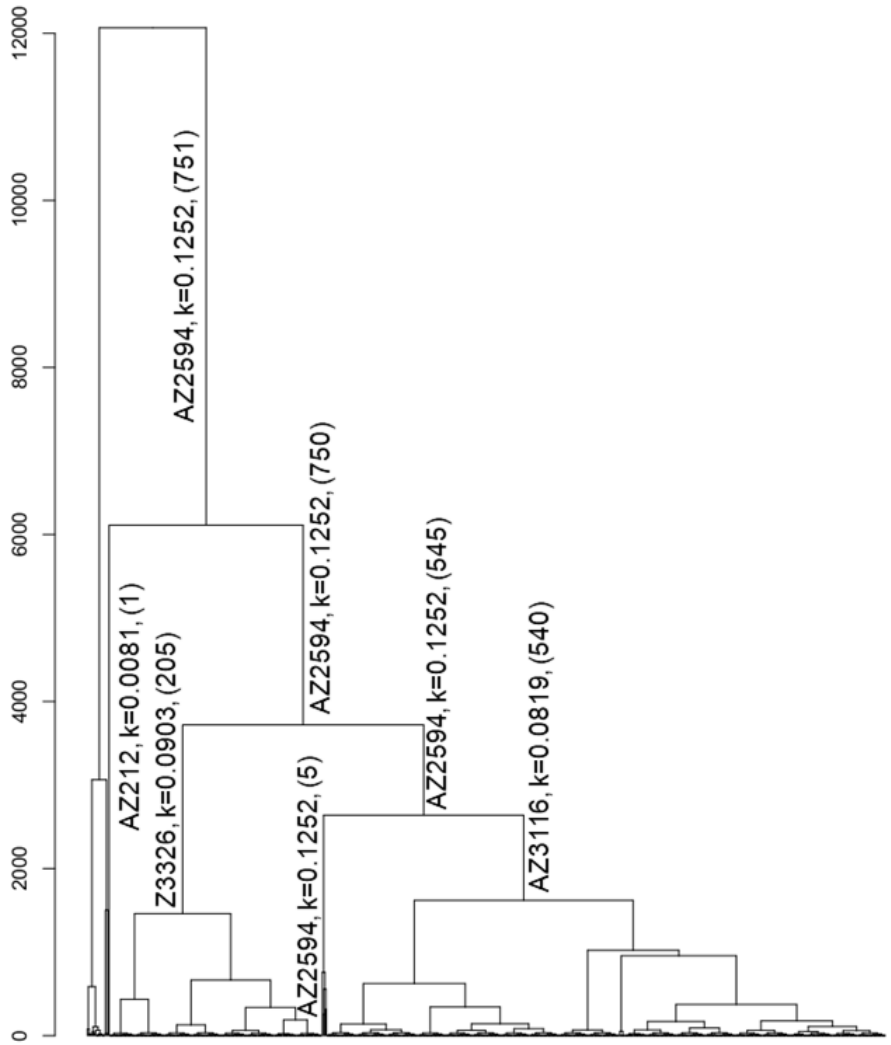

Figure 5.2 Dendrogram drawn based on a distance matrix of *T. ambiguum*. Accessions indicated on the fusion points are influential parents determined by highest average kinship (k) in the clade. The number of accessions within clades are indicated in parentheses.

Trifolium arvense

There were four accessions that strongly influenced the population structure of *T. arvense* (Figure 5.3); AZ2252, AZ6228, AZ4764 and AZ6662. Accession AZ2252 (*k*=0.025) was listed as an accession in 1986 and had the half parentage of AZ1353. AZ6228 (*k*=0.007) was listed as an accession in 2013 and had the half parentage of AZ4755. Accession AZ4764 (*k*=0.028) was listed as an accession in 2002 and had the half parentage of AZ2252. Accession AZ6662 (*k*=0.012) was listed as an accession in 2014 and had the half parentage of AZ2925.

Clusters 1, 3 and 4 of the dendrogram (Figure 5.3) can be explained by country of origin. Clusters 1, 3 and 4 had accessions with specific country of origin. Cluster 1 had Turkey (11), cluster 4 had Turkey (1), New Zealand (5) and Romania (1). Cluster 3 had Cyprus (3), Armenia (3), Yugoslavia (1) and Greece (9).

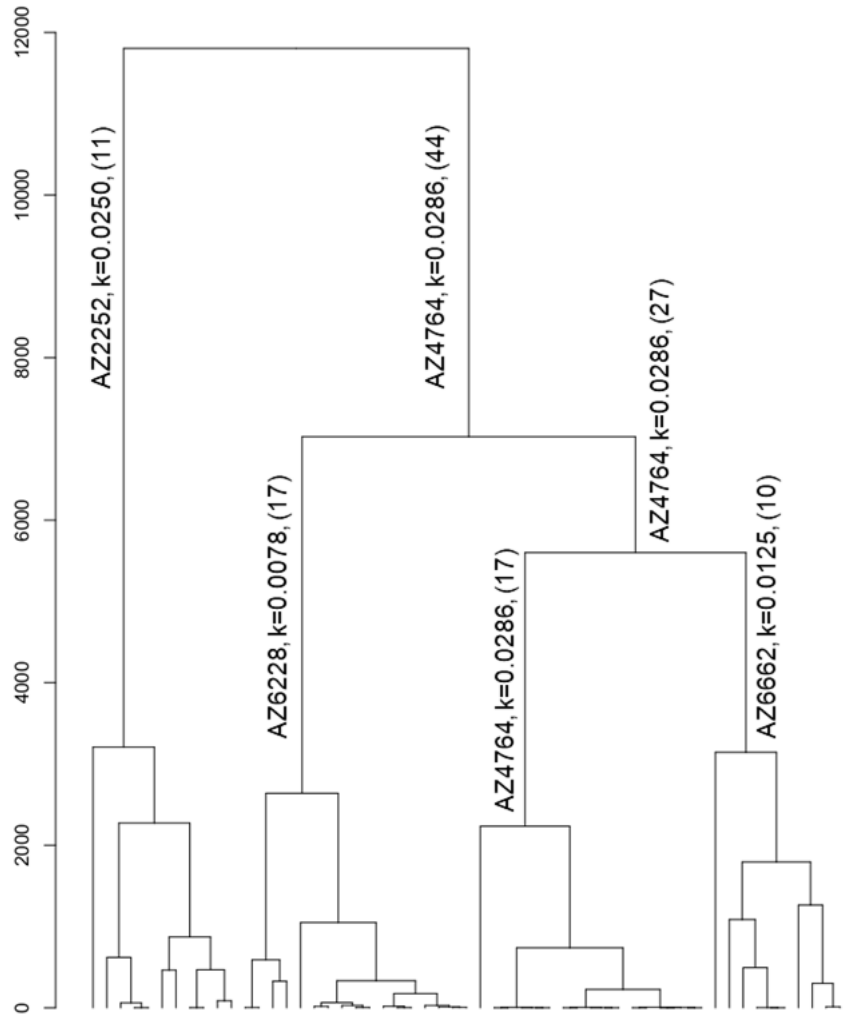

Figure 5.3 Dendrogram drawn based on a distance matrix of *T. arvense*. Accessions indicated on the fusion points are influential parents determined by highest average kinship (k) in the clade. The number of accessions within clades are indicated in parentheses.

Trifolium dubium

There were three accessions that strongly influenced the population structure of *T. dubium* (Figure 5.4); AZ4562, AZ2546 and AZ2019. Accession AZ4562 (*k*=0.001) was listed as an accession in 1998 and was known as 'suckling clover'; the parentage listed was AZ2641. Suckling clover is native to Europe and has been found to grow in subarctic regions (Nawrocki, 2011). Accession AZ2546 (*k*=0.038) was listed as an accession in 1987 and had the half parentage of AZ2019. Accession AZ2019 (*k*=0.029) was listed as an accession in 1986 and had the half parentage of AZ753.

Clusters 3 and 4 in the dendrogram (Figure 5.4) can be explained by country of origin. Cluster 4 had three countries with one accession from, New Zealand, France and Portugal. Cluster 3 had Romania (2), New Zealand (2), France (1) and the UK (1).

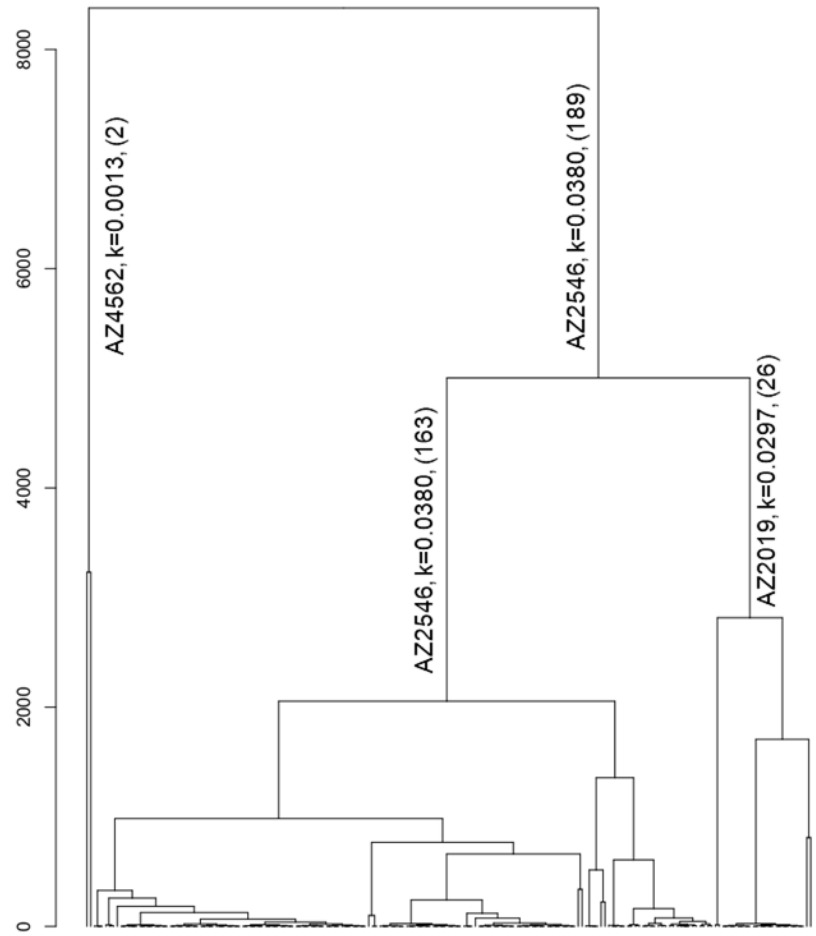

Figure 5.4 Dendrogram drawn based on a distance matrix of *T. dubium*. Accessions indicated on the fusion points are influential parents determined by highest average kinship (k) in the clade. The number of accessions within clades are indicated in parentheses.

Trifolium hybridum

There were four accessions that strongly influenced the population structure of *T. hybridum* (Figure 5.5); AB444, AB402, AB290 and AB275. Accession AB444 (*k*=0.004) was listed as an accession in 1991 and had AB402 as half parentage listed. Accession AB402 (*k*=0.025) was listed as an accession in September 1988 and had the half parentage listed as AB290. Accession AB290 (*k*=0.011) was listed as an accession in 1984 and had the half parentage listed as AB230. Accession AB275 (*k*=0.003546) was listed as an accession in 1984 and had the half parentage of AB75 listed.

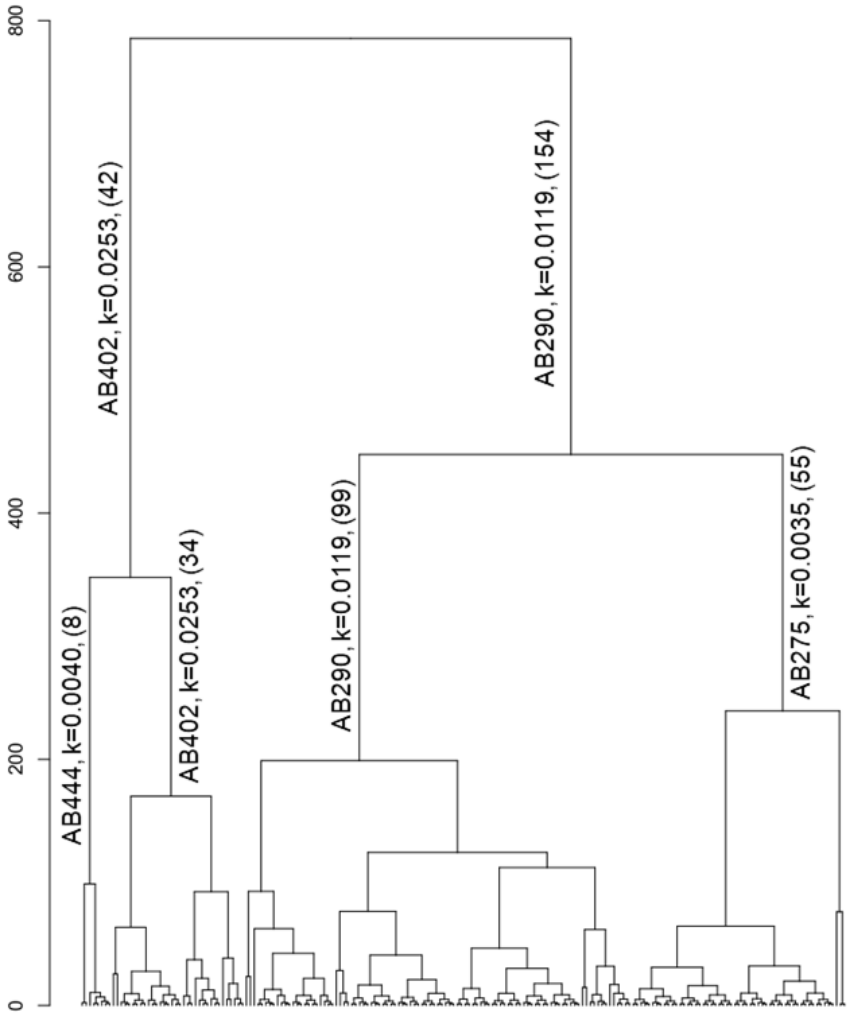

Figure 5.5 Dendrogram drawn based on a distance matrix of *T. hybridum*. Accessions indicated on the fusion points are influential parents determined by highest average kinship (k) in the clade. The number of accessions within clades are indicated in parentheses.

Trifolium medium

There were four accessions that strongly influenced *T. medium* (Figure 5.6); Z79, Z73, Z234 and Z157. Accession Z79 (*k*=0.063) was listed as an accession in 1954 and had the full parentage listed as Z13/Z16. Accession Z73 (*k*=0.061) was listed as an accession in 1954 and had the full parentage of Z6/Z16. Accession Z234 (*k*=0.012) was listed as an accession in 2014 and had the half parentage of Z13. Accession Z157 (*k*=0.125) was listed as an accession in 1989 and had the half parentage of Z127.

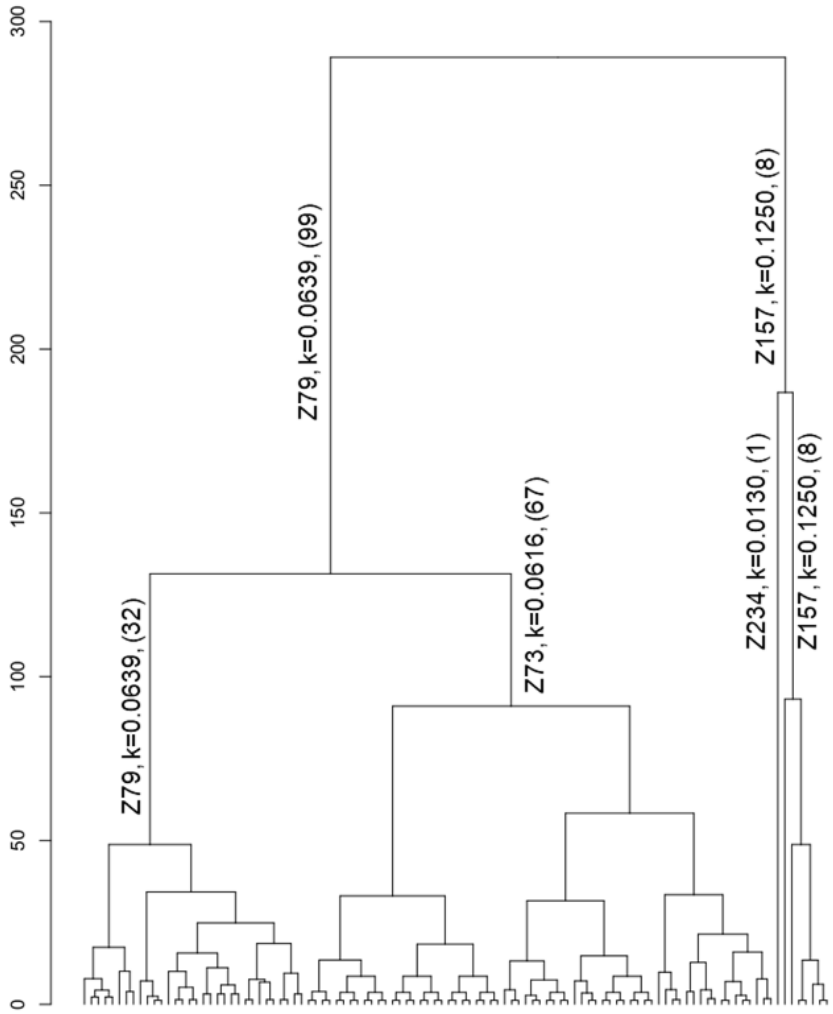

Figure 5.6 Dendrogram drawn based on a distance matrix of *T. medium*. Accessions indicated on the fusion points are influential parents determined by highest average kinship (k) in the clade. The number of accessions within clades are indicated in parentheses.

Trifolium subterraneum

There were four accessions that strongly influenced *T. subterraneum* (Figure 5.7); AK871, AK982, AK1309 and AK1333. Accession AK871 (*k*=0.002) was listed as an accession in 1988 and had the half parentage of AK576. Accession AK982 (*k*=0.009) was listed as an accession in 1989 and had the half parentage of AK799. Accession AK1309 (*k*=0.009) was listed as an accession in 2014 and had the half parentage of AK1213. Accession AK1333 (*k*=0.001) was listed as an accession in 2015 and had the half parentage of AK982.

Identifiable proportions of the dendrogram (Figure 5.7) can be explained by the country of origin. Cluster 2 has 64 accessions that are associated with a specific country; Morocco (5), New Zealand (7) Algeria (2), France (1), Greece (22), Israel (2), Italy (4), Portugal (12), Spain (6), Tunisia (1) and Yugoslavia (2). In comparison, cluster 4 had 15 accessions associated with a listed country; France (1), New Zealand (2), Italy (4) and Portugal (8).

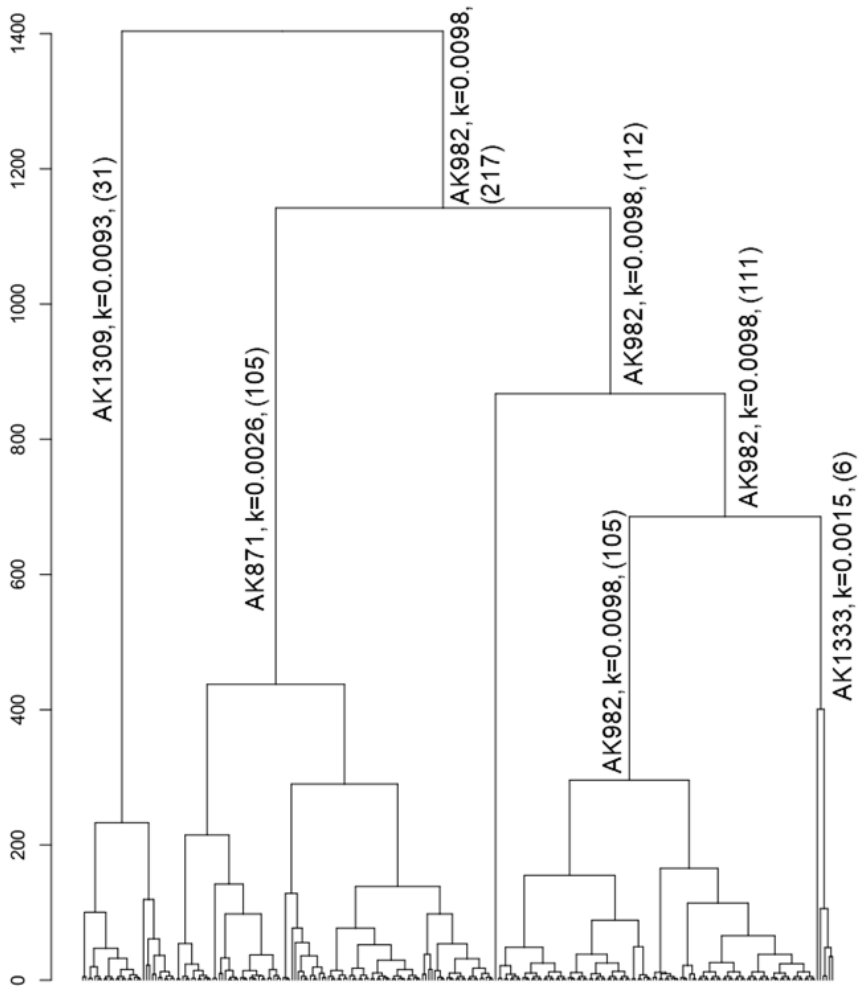

Figure 5.7 Dendrogram drawn based on a distance matrix of *T. subterraneum*. Accessions indicated on the fusion points are influential parents determined by highest average kinship (k) in the clade. The number of accessions within clades are indicated in parentheses.

T. repens x *T. occidentale* ISH

There were three accessions that strongly influenced *T. repens x T. occidentale* (Figure 5.8); AZH446, AZH1605 and AZH1884. Accession AZH446 (*k*=0.001) was listed as an accession in 2016 and had the half parentage listed of C25897 and a commercial white clover cultivar, 'Trophy'. Accession AZH1605 (*k*=0.008) was listed as an accession in 2016 and had the half parentage of AZH784. Accession AZH1884 (*k*=0.008) was listed as an accession in 2016 and had the full parentage listed of AZH1608/AZH1605. The accession AZH1605 had the ancestor of Kopu II and Durana, and AZH1884 had the ancestor Durana. Both Kopu II and Durana are commonly used commercial varieties in the New Zealand pastoral sector.

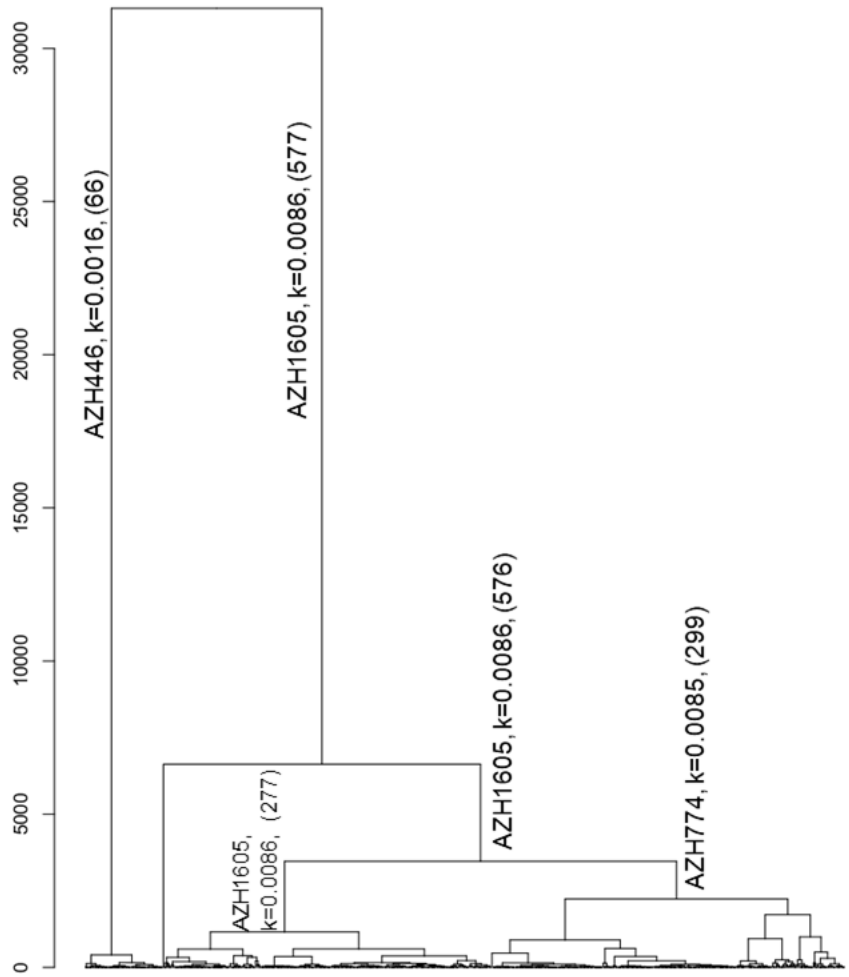

Figure 5.8 Dendrogram drawn based on a distance matrix of *T. repens* x *T. occidentale* interspecific hybrids (g). Accessions indicated on the fusion points are influential parents determined by highest average kinship (k) in the clade. The number of accessions within clades are indicated in parentheses.

5.3.4 Half kinships, indirect relationships and unrelated accessions

A total of 298,378 pairwise combinations were calculated within *T. ambiguum*, from 772 accessions. The overall kinship mean was 0.009, representing the overall average diversity of 0.99 based on parentage information. Out of the total 298,378 pairwise combinations, 289,775 (97.12%) had *k*=0, 2,581 (0.86%) had *k*=0.0625, 2,904 (0.97%) had *k*=0.125, 2,197 (0.74%) had *k*=0.25 and 149 (0.04%) had *k*=0.5.

A total of 1,540 pairwise combinations were calculated within *T. arvense*, from 55 accessions. The overall kinship mean was 0.004, representing the overall average diversity of 0.996 based on parentage information. Out of the total 1,540 pairwise combinations, 1,467 (95.26%) had *k*=00, 10 (0.65%) had *k*=0.25 and 8 (0.52%) had *k*=0.5.

A total of 18,336 pairwise combinations were calculated within *T. dubium*, from 191 accessions. The overall kinship mean was 0.005, representing the overall average diversity of 0.995 based on parentage information. Out of the total 18,336 pairwise combinations, 17,790 (97.02%) had k=0, 15 (0.08%) had k=0.0625, 108 (0.59%) had k=0.125, 182 (0.99%) had k=0.25 and 49 (0.27%) had k=0.5.

A total of 19,110 pairwise combinations were calculated within *T. hybridum*, from 196 accessions. The overall kinship mean was 0.001, representing the overall average diversity of 0.999 based on parentage information. Out of the total 19,110 pairwise combinations 19,048 (99.68%) had k=0, 8 (0.04%) had k=0.125, 18 (0.09%) had k=0.25 and 36 (0.19%) had k=0.5.

A total of 5,778 pairwise combinations were calculated within *T. medium*, from 107 accessions. The overall kinship mean was 0.022, representing the overall average diversity of 0.978 based on parentage information. Out of the total 5,778 pairwise combinations, 5,260 (91.03%) had k=0, 332 (5.75%) had k=0.25 and 79 (1.37%) had k=0.5.

A total of 30,876 pairwise combinations were calculated within *T. subterraneum*, from 248 accessions. The overall kinship mean was 0.0002, representing the overall average diversity of 0.9998. Out of the total 30,876 pairwise combinations 30,613 (99.15%) had k=0, 1 (0.003%) had k=0.25 and 14 (0.05%) had k=0.5.

A total of 207,046 pairwise combinations were calculated within *T. repens* x *T. occidentale*. The overall kinship mean value was 0.004, representing the overall average diversity of 0.996 based on parentage information. Out of a total 207,046 pairwise combinations, 206,126 (99.56%) had k=0, 4 (0.002%) had k=0.13, 141 (0.07%) had k=0.25, 132 (0.06%) had k=0.5.

5.4 Discussion

5.4.1 *T. ambiguum* pedigree map complexity

The lineage of the cultivar Monaro had a large contribution to the pedigree of *T. ambiguum*. Monaro is a hexaploid (6x=48) Australian cultivar described as vigorous and productive. It is suitable for most environments, especially areas where there is periodic summer drought (Fu, et al., 2001, Oram, 1977, Virgona, et al., 1996). Virgona, et al. (1996) examined the performance of Monaro 11 years after establishment in south-eastern Australia against *T. repens* and *T. subterraneum*. Monaro was superior in legume content, digestibility and ability to respond to additional fertiliser.

The large root and rhizome systems of *T. ambiguum* and its well described tolerance to drought (Oram, 1977, Sheaffer, et al., 2009, Virgona, et al., 1996) has prompted its crossing with white clover (*T. repens* x *T. ambiguum*). White clover has a shallow root system which leads to underperformance in drought conditions. The *T. repens* x *T. ambiguum* ISH breeding programme has shown success in integrating larger root systems into the hybrid progeny (Abberton, et al., 1998, Marshall, et al., 2015). Abberton, et al. (1998) showed that BC3 hybrids were outperforming the other hybrids, *T. ambiguum*

and *T. repens* in root characteristics. Marshall, et al. (2003) showed that comparisons of above-and-below ground biomass for BC1 and BC2 hybrid plots showed more dry matter in roots and rhizome of clover than in the *T. repens* plot.

5.4.2 Influencing accessions and pre-breeding traits

Over the period (1950-2016) there have been many international introductions of *Trifolium* species into the MFGC (Figure 5.1). The introduction of germplasm from other countries has increased the variation in the MFGC's clover collection and is crucial for maintaining the improvement of elite clover material. Introduced germplasm is important in pre-breeding programmes and often become the parents that influence the population structure.

T. subterraneum is one of Australia's most important forage legumes and has adapted to the harsh Mediterranean climates in specific parts of Australia (Nichols, et al., 2007a, Nichols, et al., 2013b). Introductions from the semi-arid conditions of South Australia and Morocco can provide insight into the adaptation extremes of subterranean clover. Having the adaptation to dry climates can be beneficial for drought-tolerant breeding programmes. In New Zealand, this is a prominent expectation of subterranean clover and is being trialled on dry areas of both hill and high-country farms to replace the use of white clover where irrigation is not available (Lucas, et al., 2015, Macfarlane, et al., 1990a, Macfarlane, et al., 1984, Widdup, et al., 2000). Using more drought-tolerant *T. subterraneum* will increase its versatility in location of sowing and ability to withstand and persist in dry conditions.

The *T. repens* x *T. occidentale* ISH had introductions from countries that have dry, arid conditions, suitable for drought-tolerant phenotypes. The common objective of breeding programmes for *T. subterraneum* and *T. repens* x *T. occidentale* is to introgress drought tolerance traits, often through breeding for persistence in the former and larger root systems in the latter.

Kopu II was present in the pedigree of the *T. repens* x *T. occidentale* ISH. Kopu II is a large leaf, highly persistent and yielding white clover that has rapid recovery after grazing and elevated sugar levels. A crucial advantage of Kopu II is the high tolerance to clover root weevil, which is becoming an increasing problem in pastoral systems (Woodfield et al., 2001). Durana was bred specifically to be highly persistent under grazing pressure and to tolerate acidic soils (Bouton, et al., 2005). The commonality of the two commercial cultivars having persistence as the key objective could be what is influencing the ancestors to be at the top of the clusters. While white clover has suitable persistence for multiple pastoral systems, *T. occidentale* is not as broadly adapted. By using the cultivars that are very high in persistence in an ISH, it could reduce the risk of losing the persistence traits in recombination events.

Climatic and geographical similarities between the influencing accessions in all *Trifolium* spp is strong reasoning behind the divergence and influence of the founding accessions. The regions and countries associated with accessions provide insight into desirable phenotypic advantages. There were several

accessions originating from Australia, which has several climate zones due to the latitudinal span of the country, ranging from temperate to tropical climates. However, the largest area of Australia is desert and semi-arid. Therefore, the phenotypes of plant introductions from Australia are often drought tolerant and adapted to arid climates. Accessions originating from the Mediterranean region will have similar adaptations to Australia due to similar climates (Bolle, 2003). The Mediterranean has dry summers and rainy winters, often producing collected ecotype accessions with drought-tolerant phenotypes.

There were three regions in New Zealand where accessions have originated from: Manawatu, Hawkes Bay and Mackenzie country. The Manawatu region in New Zealand is not exposed to extreme climatic conditions. However, drought conditions can occur in summer (Chappell, 2015). The Hawkes Bay region is prone to rainfall and temperature variations. The summer season is drought-prone with prevailing winds (Thompson, 1987). The Mackenzie country is one of the driest areas of New Zealand. The region has a temperate-continental climate with long, hot summers and snow in winter (NIWA, 2017). Within the *Trifolium* spp. introductions, drought-tolerant traits are often a desirable adaptation to have and are a common breeding objective. As the species are used in pastoral systems in New Zealand, both irrigated and non-irrigated, it is often desirable that the species have summer drought tolerance.

5.4.3 Diversity of *Trifolium* accessions

The kinship values indicate that genetic relatedness within *Trifolium* spp. is low (Figure 5.9). However, kinship, as derived in this study, overestimates genetic diversity because we assume that all relevant parents are unrelated, as they are giving rise to different clusters in the population structure. However, this does not have to be true. Crucial genes that control species-specific characteristics (e.g., leaf size and trifoliate leaf) must be homologous so that the individual can be classified as a member of the species (Graczyk, et al., 2015). Kinship may overestimate genetic diversity because of unaccounted relationships due to the lack of bookkeeping when recording crosses. In all *Trifolium* spp. reported here, there were accessions with null parentage listed (Table 5.1). These were excluded from the genetic parameter analysis due to the questionable results of whether a zero kinship was from null parentage or a true zero k value. The correlation of estimates of genetic diversity based on pedigree data have been reported for other plant species (Lacombe, et al., 2013, Navabi, et al., 2014). However, missing data in the parental information can influence correlations to be poor, with higher correlations being expected when the available pedigree information is more detailed (Lacombe, et al., 2013). El-Kassaby, et al. (2011) argues that when looking at a pedigree on a species level, not all crosses need to be accurately recorded as the pedigree is producing a large overview of population-specific patterns, rather than within cluster or accession analysis. If the purpose of the study is to analyse within cluster or accession patterns, it is crucial that the pedigree data is complete or near-complete.

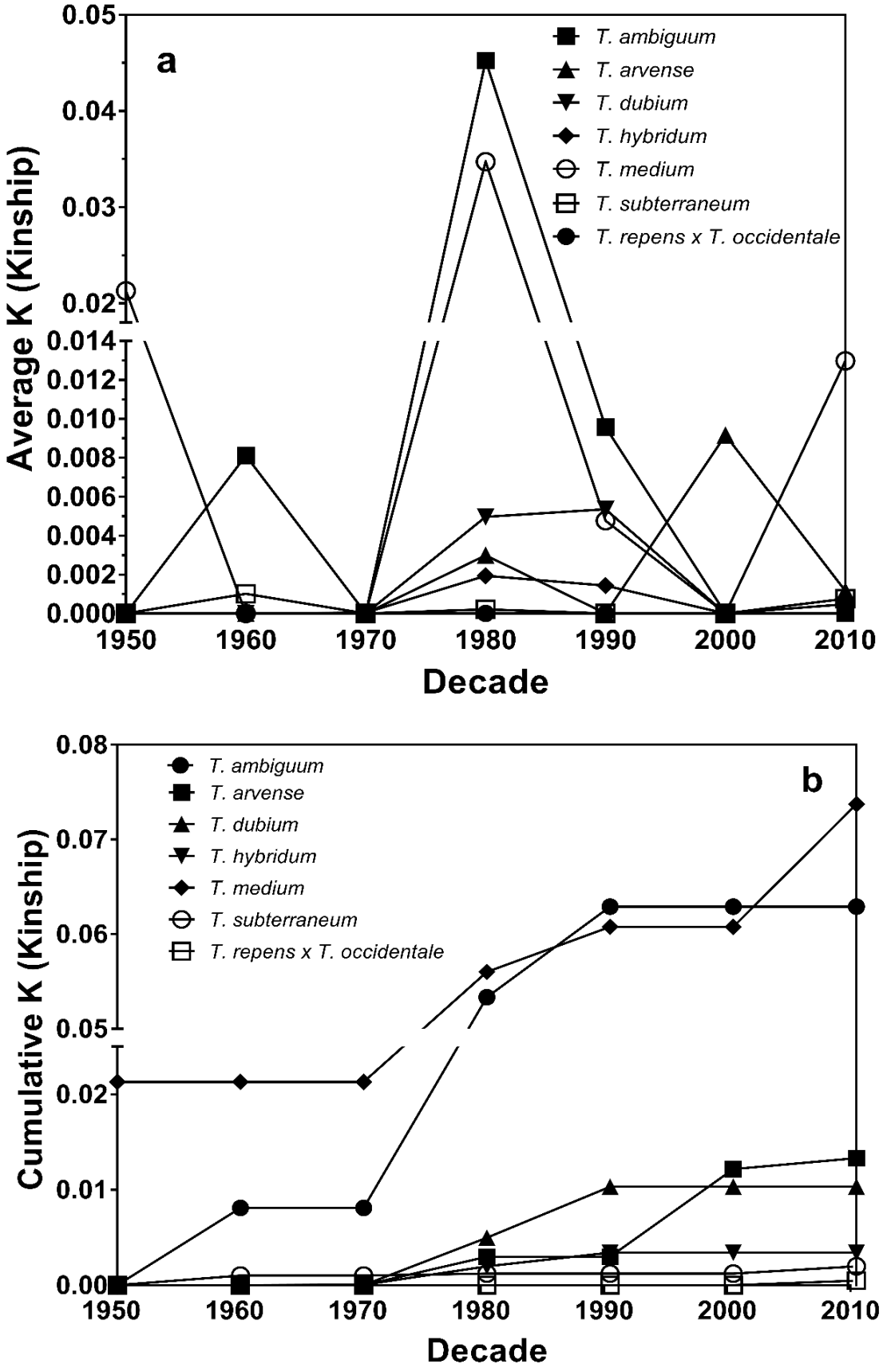

Figure 5.9 The trend in average (a) and cumulative (b) kinship in seven *Trifolium* species across seven decades.

Interestingly, a large proportion of accessions in *Trifolium* spp. had an average kinship coefficient of 0, showing that they are completely unrelated to any accession within the species. This indicates useful crosses that could be planned across breeding pools to generate new variation. Populations with a wider genetic base and greater adaptation could produce offspring with important agronomic traits. Within species diversity is a topic of high importance to many pastoral research groups because of the documented benefits of greater genetic variation (Sanderson, et al., 2004).

Intra-population genetic diversity is important for the long-term persistence of breeding programmes for two main reasons: 1) phenotypic variation is correlated with adaptive potential of populations, and 2) neutral genetic variation of natural populations reflects inbreeding and genetic drift, which reduces the variation within populations (Blows, et al., 2005, Kahilainen, et al., 2014, Reed, et al., 2003).

5.4.4 Genetic resources for *Trifolium* improvement

The *Trifolium* genus is held in many genebanks globally, comprising of landraces, accessions, breeding lines, wild relatives and commercial cultivars (Abbasi, 2008, Morris, et al., 2001, Rosso, et al., 2005, Rosso, et al., 2001). However, it has been noted that often *Trifolium* collections are biased towards the two major species, red and white clover. The first clover entry into the MFGC was in 1931 and 1897 in the United States collection. In the USDA collection red and white clover contribute to 56% of the collection. Wild relatives are present; however, the presence is poor for species that are possible gene sources. Only 4.5% of the collection is species that are close relatives to red and white clover (Morris, et al., 2001).

The results from the *Trifolium* species in this study are comparable to the results by Egan, et al. (2019a) and Egan, et al. (2019b), where diversity within the white and red clover accessions at the MFGC were assessed. Overall, the level of variation was high across all of the *Trifolium* species accessions. These results are encouraging as the mating systems of *Trifolium* species are outcrossing and self-incompatible. Utilising the available variation within the accessions will assist breeders in developing populations for specific breeding targets. While it is useful to assess variation within each species, it is optimal to compare the results across species to harness all available variation for pre-breeding and hybridisation programmes.

The high levels of diversity found in the genebank collections reported in white (Egan, et al., 2019a) and red clover (Egan, et al., 2019b) are positive for future hybridisation programmes. Pre-breeding efforts in minor *Trifolium* species could complement white and red clover breeding. Egan, et al. (2019a) showed that the country of origin of accessions has influence on the population structure. Hybridisation programmes of major *Trifolium* species with characterised germplasm of minor species could utilise that information and increase the efficiency of crosses. The now-defined breeding pools in this study could be exploited for between- and-within pool crosses with white and red clover.

5.4.5 Pre-breeding from related species for *Trifolium* improvement

Pre-breeding is an essential bridge between genebanks and breeding programs. Pre-breeding is defined as all activities designed to identify useful characteristics from unadapted germplasm. The goals of pre-breeding can be broadly classified as widening the base of diversity and to increase plant production through various traits. The programs developed from pre-breeding can generate new base populations with the end goal of cultivar development. When used in hybrid programs, heterotic patterns can be identified.

The most prominent example of pre-breeding success in *Trifolium* is the development of the white clover ISH (Williams, 2014). Williams, et al. (2006) identified the 10 *Trifolium* species that could be used in hybridisation with white clover. It was thought that these species could bring useful traits to white clover as the minor species often have useful genes that can contribute to increased genetic gain (Abberton, 2007, Williams, 2010). Hoyos-Villegas, et al. (2019) investigated the rate of genetic gain in white clover in a selection of 80 white clover cultivars released between 1920 and 2010 across 17 countries. Their study showed that there has been less than 0.17% gain per decade in both yield and content. The three species that have been popularly bred as hybrids are *Trifolium uniflorum*, *Trifolium occidentale* and *Trifolium ambiguum*. *Trifolium uniflorum* have been successful in breeding programmes and have shown promising results in producing hybrids that have increased performance under drought tolerance and reduced phosphate levels (Hussain, et al., 2017, Nichols, et al., 2014c, Nichols, et al., 2016, Nichols, et al., 2015).

There has been less research performed for *Trifolium pratense* ISH, however, *Trifolium medium* has been popular for integrating perenniality into red clover (Abberton, 2007, Merker, 1984, Sawai, et al., 1990, Sawai, et al., 1995). Isobe, et al. (2002) created four backcross generations of *T. pratense* x *T. medium* ISH. The BC_4 plants had a 61% survival rate by the fourth year of the field trial and almost all BC_4 plants produced mature seed. Female fertility was 21.3% and pollen fertility was 65.3%. These results showed the potential for these hybrids to be used in future red clover selection programs.

Challenges in pre-breeding

The ability for the germplasm held in the genebanks to respond to differing abiotic and biotic stresses is well recognised. However, as promising as pre-breeding is, there are big challenges too. The two biggest challenges are time and reducing linkage drag. A large time commitment is needed to identify useful germplasm and hybridise it with well-adapted germplasm whilst reducing the unwanted genes (Singh, et al., 2018). The time commitment is also largely dependent on the information on the accessions at the beginning of the pre-breeding programme (Nass, et al., 2000, Williams, et al., 2012). The ability to determine whether germplasm is compatible is also time consuming, especially if there is lack of knowledge around wild species. Linkage drag is a major limiting factor in pre-breeding and Sharma, et al. (2013) states that it is the most important factor responsible for low use of germplasm. Linkage drag is defined as the unplanned inheritance of undesirable genes along with the target genes

due to their close linkage (Acquaah, 2012). Generating large populations and utilising genomic tools can help to overcome linkage drag (Armstead, et al., 2001, Sharma, 2017, Sharma, et al., 2013).

5.5 Conclusions

We found that the genetic resources held in the *Trifolium* spp. at the MFGC have a wide genetic base. The pedigree maps constructed, and derived relatedness parameters estimated showed that seven *Trifolium* spp. in the MFGC have high diversity within the recorded germplasm. The absence of inbreeding in all species highlights the available genetic diversity and is a positive insight into what forage breeding has achieved. There were no visual bottlenecks in the pedigree maps.

Germplasm exchange between countries and domestic and international collection trips has proved successful in creating a collection of accessions of *Trifolium* spp. for the MFGC. Influencing ancestors have been identified and relevant parents that influenced population structure have been distinguished. The results from interrogating the pedigrees showed that geographical origin seemed to be influencing the international introductions. Knowing the country of origin of accessions that contributed large numbers of progeny and resulted in elite material, and obtaining the breeding pools to which they belong, can provide knowledge on diverging phenotypic characteristics. This will enhance future pre-breeding decisions (Zamir, 2001). Introductions from countries with semi-arid environments strongly influenced the population structures and adaptation to target environments drove divergence between the clusters of accessions.

Overall, the results from this research on the population structure of *Trifolium* spp. are relevant to pre-breeding efforts. Although *Trifolium* spp. are not as largely used as *T. repens* and *T. pratense* in agricultural systems, climate change and societal demands for sustainable agriculture will require that new traits are integrated into *T. repens* and *T. pratense* via interspecific hybridization. Species such as *T. ambiguum* can hybridise with white clover, and *T. medium* with red clover, to improve production and survival, particularly in marginal environments. The *T. repens* x *T. occidentale* ISH has been successful in the improvement of white clover root systems. With a known population structure, the future pre-breeding decisions can be more efficient and defined breeding pools will maximize germplasm utilisation. Breeders can make use of the assembled pedigree maps reported here to strategize how to maximize genetic variation and incorporate pre-breeding efforts in breeding programs. Despite the challenges associated with pre-breeding, it is critical that time and funding is invested into pre-breeding programmes. Utilising wild relatives will strengthen breeding programmes and accelerate the rate of genetic gain and the release of cultivars to market.

Chapter 6
A genome-wide association study of herbage yield and stolon density in white clover

6.1 Introduction

Herbage yield and persistence are two important traits that determine the performance of white clover. Herbage yield, often indirectly measured through a growth score in the field, is directly related to animal live-weight gain and economic profit (Chapman, 1983). Persistence is an essential trait for animal grazing. The persistence of a white clover plant is dependent on many interacting factors, including plant morphology, biochemistry and environmental conditions (Baker, et al., 1987). Persistence begins with a robust and dense network of stolons to allow the plant to recover from grazing but also persist within a sward. Persistence can be measured indirectly through a stolon density score in the field (Widdup, et al., 2011). Woodfield, et al. (1996) summarised that plant breeders could increase persistence by increasing the rate of stolon formation. However, the negative correlation between herbage yield and stolon density is difficult for breeders to develop improved populations (Jahufer, et al., 1999).

The different morphological growth stages of white clover are a challenge for breeders when trying to identify breeding lines that consistently have superior performance. It has become essential for white clover field trials to have at least three years of assessment to account for the variation in growth stages (Widdup, et al., 2011). Brock, et al. (2001) summaries the three distinct growth stages of white clover from seed as (i) seedling phase, (ii) tap-rooted phase, and (iii) clonal phase. The seedling stage has few stolon branches which are then rapidly increased in elongation in the tap-rooted phase. Due to the rapid increase in growth, the tap-rooted phase usually produces the most significant herbage yield. A mature clonal set of plants are present in the clonal phase, and the stolons are reliant on the nodal roots. Over time the larger plants are fragmented into smaller plants (Brock, et al., 2000). When white clover is in the clonal phase, herbage yield is often variable, but there is a general trend of decline (Caradus, et al., 1989c).

Widdup, et al. (2011b) states that ten years of evaluation data on five different sites is needed for a white clover cultivar to go to market. This amount of time is deemed sufficient to test for all morphological growth phases and to assess tolerance to biotic and abiotic stresses. Although this is thorough and has been successful, there is ever demanding need on increasing the rate of selection cycles.

Molecular markers have allowed the genetic dissection of complex traits and aided in selection cycles through marker-assisted selection. Next-generation sequencing (NGS) has enabled the identification of

thousands of SNPs at a low cost (Kumar, et al., 2012). However, the use of markers in white clover has been challenging. The breeding system and ploidy level complicate the genotyping and analysis of white clover, as there is variation between and within subgenomes (Kaur et al., 2012).

The deployment of genomic techniques for selection of quantitative traits has practical relevance for open-pollinated forage legumes (Biazzi, et al., 2017). The integration of molecular tools into white clover breeding programmes to increase the speed and accuracy of targeting agronomic traits is desirable. Although the development of molecular tools in forage species has proved challenging, it has been successful (Faville, et al., 2012). The development of linkage maps in white clover has aided in identifying regions of the genome that contain genes controlling traits (Zhang, et al., 2007). QTL have been identified in white clover for a variety of agronomic traits, most notably seed yield (Barrett, et al., 2005, Barrett, et al., 2009, Cogan, et al., 2006, Zhang, et al., 2010).

This chapter focuses on addressing the knowledge gap in understanding the genetic architecture of complex traits in white clover. The objectives of this chapter were to evaluate a panel of 242 white clover genotypes to (i) investigate population structure, and (ii) identify genomic regions that control herbage yield and stolon density.

6.2 Materials and Methods

6.2.1 Plant material and phenotyping

The plant material used in this study was 242 white clover half-sib families used in a development population for the cultivar Mainstay. The population was selected for GWAS as it was diverse material being developed into elite populations, that had been phenotyped as part of a trial in a previous programme. The half-sib families were crossed in a polycross cage, and the seed was harvested from each maternal plant. 250 half-sib families were originally included in the population, but the number was reduced to 242 due to low numbers of seed for some of the half-sib families. The genotypes were phenotyped in Palmerston North, New Zealand from the 2011 summer season to the 2014 autumn season. The growth score was scored on a scale of 0-10 with 0 being low growth and 10 being high growth. The stolon density was scored on a range of 1-5 with 1 being low density and 5 being high density. The parental germplasm originated from crosses between the white clover cultivars, Kopu II and Barblanca (Crush, et al., 2015).

6.2.2 Genetic markers

DNA preparation and sequencing

The plant material was leaf material that had been freeze-dried for five years. The DNA extraction methodology was the same as described in Anderson, et al. (2018). Briefly, the plant material (15-24 individuals per half-sib family) were ground with a bead beater until the leaf material was in a powder form. The homogenisation buffer was added and centrifuged. The supernatant was removed, and the

precipitation buffer was added before mixing and incubating the plates. The plates were centrifuged at maximum speed for 30 minutes at room temperature. A binding buffer and the supernatant were mixed and centrifuged. The plates were then washed with a binding buffer and a wash buffer and centrifuged after each buffer was added. The plates were washed with ethanol and centrifuged. A new plate was used to collect the DNA. Tris HCl and RNAse A were added to the plates and centrifuged. The process yielded approximately 100µl of eluent containing 1-13µg DNA per well. The DNA was quantified with a Qubit® 2.0 Fluorometer, and 230-280nm absorbance scans performed on a Nanodrop ND-1000 Spectrophotometer. Quality control of the DNA was performed by using 0.8% for volume for TBE gel per 100ml of TBE buffer. DNA was loaded into the wells and underwent electrophoresis. The gels were stained by ethidium bromide and visualised with a UV transilluminator.

The 242 white clover genotypes were genotyped using GBS. The GBS was performed through a modification of the Elshire, et al. (2011) and carried out at Cornell University. The methodology of the GBS protocol was based on Griffiths, et al. (2019) and the restriction enzyme PST1 was used. The genotypes were sequenced on an Illumina Hi-Seq.

SNP calling and filtering

Sequence reads were quality filtered and trimmed using Trimmomatic, removing sequence with an average quality below 20 on a sliding window of 10 nucleotides. Reads shorter than 100bp after trimming were discarded. The filtered reads were processed with the Tassel 5 pipeline, using the clover genome version 5 as reference. All parameters were left with default values except the minimum minor allele frequency, set to 0.001. The tags identified by Tassel5 were exported into a vcf format table. vcftools (Danecek, et al., 2011) is a program package designed for working with VCF files to provide methods for analysing complex genetic variation data in the form of VCF files. vcftools filtered the SNPs. The SNP filtering eliminated markers with a minor allele frequency (MAF) of <5%, markers with >80% missing data and depth of <5 and >300. SNPs were filtered for Hardy-Weinberg equilibrium (HWE). The final set of SNPs contained a total of 5,543 SNPs. KGD is a package for the analysis of GBS data. Primarily, KGD constructs a genomic relationship matrix for genotypes but also has quality control, relationship estimation and pedigree verification tools. KGD was used to construct a genomic relationship matrix and visualise SNP characteristics (Dodds, et al., 2015).

6.2.3 Statistical analyses

Linear mixed models

Best linear unbiased predictions (BLUPs) were used to minimise the effect of environmental and year variation for the phenotypic data to input into the GWAS model. A mixed model in GenStat was used (VSN-International, 2019):

$$y = X\beta + Zu + \varepsilon$$

where y is a known vector of observations, β is an unknown vector of fixed effects, u is an unknown vector of random errors, ε is an unknown vector of random errors, and X and Z are known design matrices relating the observations y to β and u.

Principal component analysis

The ordination technique of principal component analysis was calculated in GenStat using the sum of squares and products (VSN-International, 2019). All data were normalised by log transformation in Minitab (Minitab, 2006). Biplots visualised the variation within the population and the relationships among traits.

6.2.4 Population structure and linkage disequilibrium analysis

Plink was used to obtain the required input files (Purcell, et al., 2007). The python package fastStructure determined population structure (Raj, et al., 2014). Subpopulations were simulated for k=1 to k=10 to determine the mean of the variation posterior distribution over admixture proportions. The maximum likelihood estimation of individual ancestries from SNPs was calculated in ADMIXTURE to determine the optimal number of subpopulations (Alexander, et al., 2009). A neighbour-joining cladogram was generated in TASSEL (Bradbury, et al., 2007). The output Newick file was used for the input into the program FigTree to produce the cladogram (Rambaut, 2009).

Plink was used to calculate linkage disequilibrium (LD) decay. The following equation was used (Hill, et al., 1968, Lewontin, 1964):

$$r^2(p_a, p_b, p_{ab}) = \frac{(p_{ab} - p_a p_b)^2}{p_a(1 - p_a)p_b(1 - p_b)}$$

where p_{ab} is the frequency of haplotypes having allele a at locus 1 and allele b at locus 2.

6.2.5 Genome-wide association analysis

ASReml was used to fit a mixed model to obtain the residuals to use as phenotypes. The general linear mixed model equation used was:

$$y = X\tau + Zu + e$$

where, y (n x 1) signifies the vector of observations, τ (p x 1) is a vector of the fixed effects, X (n x p) is the matrix that associates observations with the combination of fixed effects, u (q x 1) is a vector of random effects, Z (n x q) is the matrix that associates observations with the combination of random effects, and e (n x 1) is the vector of residual errors (Gilmour, et al., 2015)

Probabilities of each genotype were calculated as in Li (2011) for each sample and SNP, assuming a sequencing error rate of 0.1%. The probabilities were used as a predictor variable in the GWAS analysis.

Marker-trait associations were analysed using the program ProbABEL (Aulchenko, et al., 2010):

$$E[\mathrm{Y}] = \mathrm{X}\beta = \mathrm{X}_x\beta_x + \mathrm{X}_g\beta_g$$

where, Y is the vector of phenotypic values, X_g is the design matrix containing data about predictors of interest, and X_x is the design matrix containing other covariates. β_g and β_x are the vectors of corresponding fixed effects. An additive model was fitted as the BLUPs were considered as breeding values. Manhattan and quantile-quantile (QQ) plots were generated with the 'manhatplot' function from GBS-PopGen on KGD GitHub (Dodds, et al., 2015). Significant SNPs were tested by recalculating the BLUPs in GenStat using the dosage variable for the SNP as a covariate.

6.3 Results

6.3.1 Genetic markers

A total of 5,543 SNPs were present after filtering. The mean and minimum co-call rate for sample pairs was 0.69 and 0.08, respectively. The mean sample depth was 48.91, and the mean call rate was 0.81. The MAF values ranged from 0.05-0.5 (Table 6.1). The SNP depth is visualised in Figure 6.1. The genomic relationship matrix generated from KGD calculated an average relatedness of 0.42.

Table 6.1 The number of markers in the specified minor allele frequency range and the corresponding average SNP depth and Hardy-Weinberg equilibrium *P*-value.

Minor allele frequency range	Number of markers	Average SNP depth	Average Hardy-Weinberg equilibrium *P*-value
0-0.1	2298	36.20	0.1986
0.11-0.2	1367	58.33	0.2059
0.21-0.3	743	63.67	0.2033
0.31-0.4	595	70.66	0.1825
0.41-0.5	540	73.05	0.1997

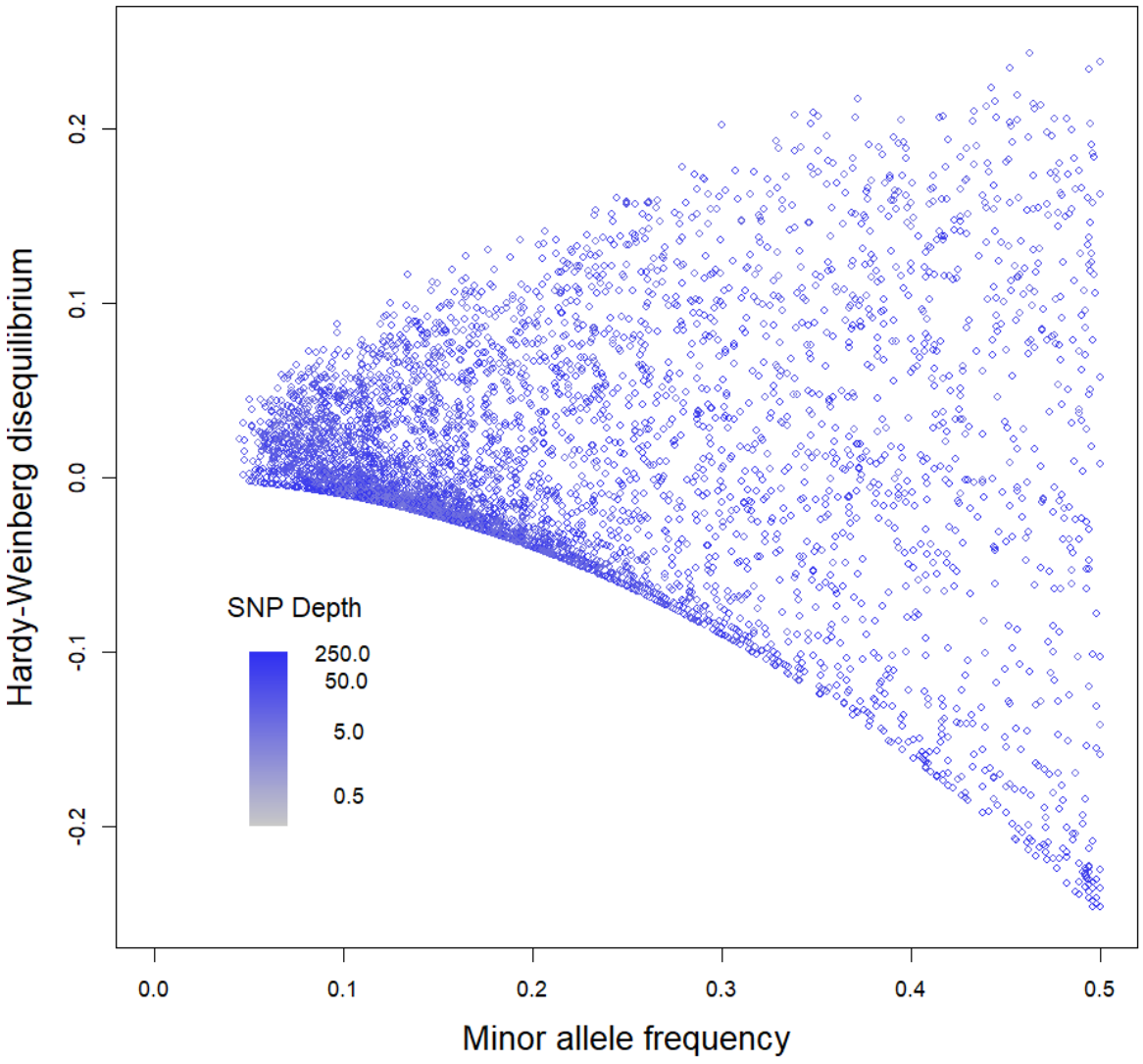

Figure 6.1 Finplot of the 5,543 filtered SNPs in the 242 white clover genotypes. The Hardy-Weinberg disequilibrium is plotted against the minor allele frequency (MAF) and shaded by the SNP depth.

6.3.2 Phenotypic data

A total of 13,598 growth scores were analysed. The mean and median growth score was 5.55 and 6, respectively. The values ranged from 0-10 (Figure 6.2). A total of 544 stolon density scores were analysed. The mean and median stolon density score was 2.77 and 3, respectively. The values ranged from 1-5 (Figure 6.3). Accessions with high growth and stolon density scores (GC216_139, GC216_272, GC216_186) and accessions with low growth and stolon density scores (GC216_285, GC216_261, GC216_004) are indicated in Figures 6.2 and 6.3.

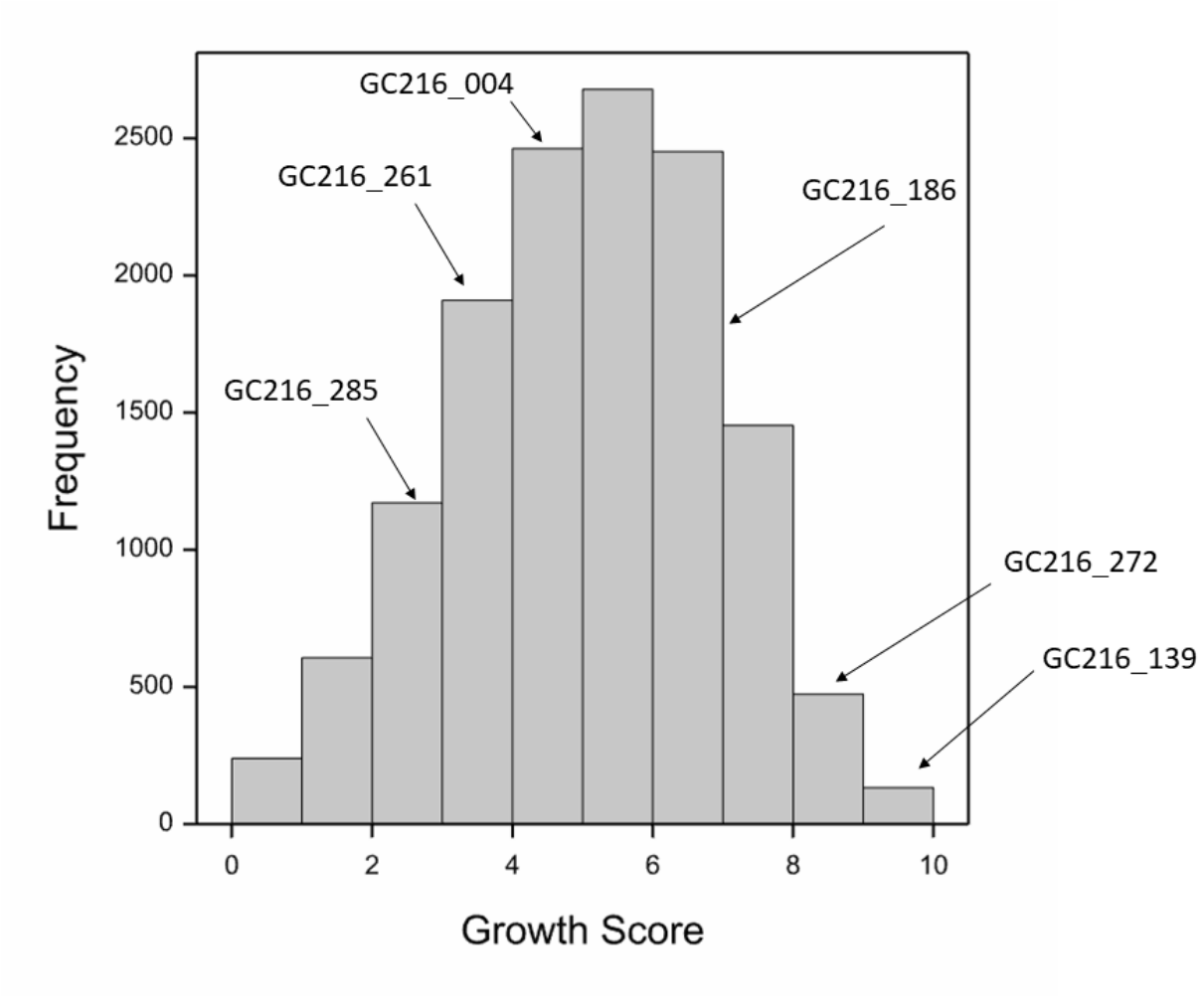

Figure 6.2 Histogram of the growth score of the 242 white clover genotypes. The bins with accessions with high growth scores (GC216_139, GC216_272, GC216_186) and low growth scores (GC216_285, GC216_261, GC216_004) are indicated.

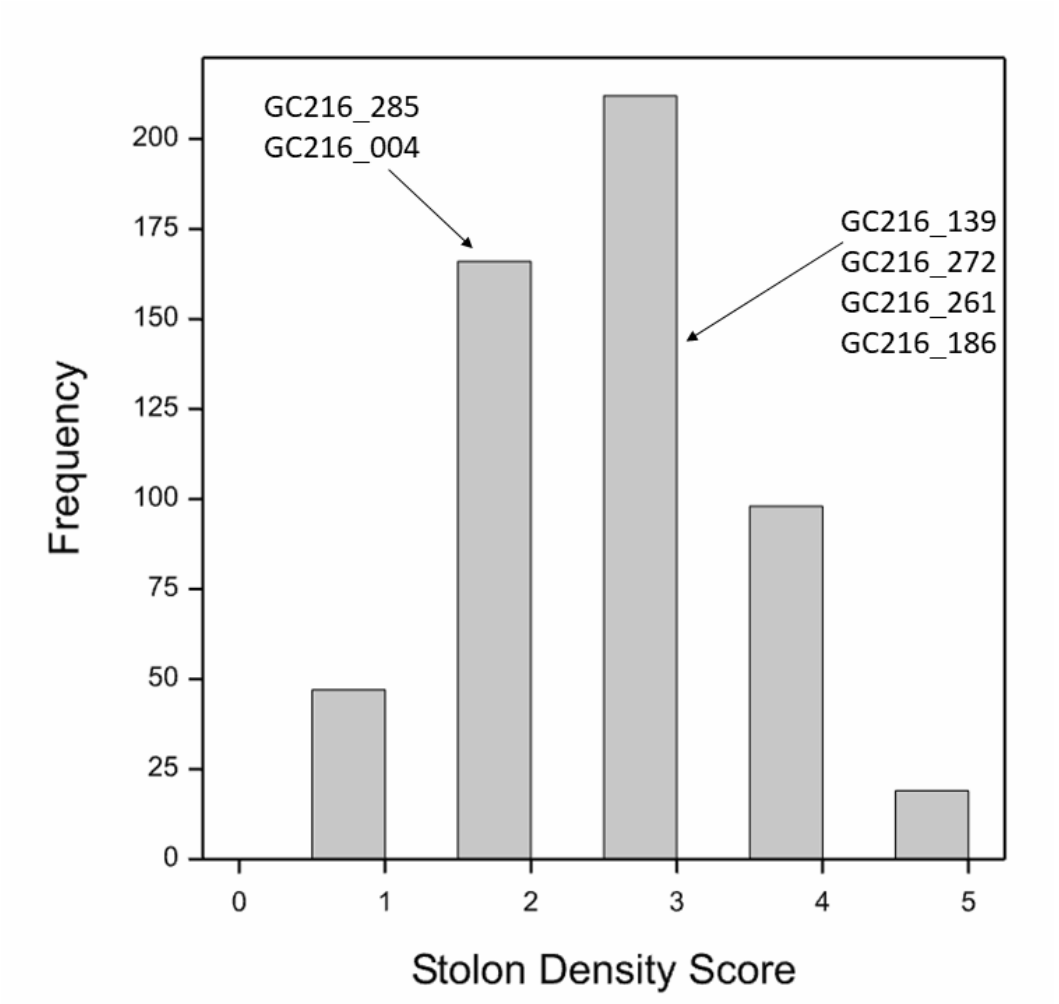

Figure 6.3 Histogram of the stolon density scores of the 242 white clover genotypes. The bins with accessions with high stolon density scores (GC216_139, GC216_272, GC216_186) and low stolon density scores (GC216_285, GC216_261, GC216_004) are indicated.

Principal component (PC) analysis of the BLUPs of overall stolon density, overall yield, years 1, 2, 3 and 4 growth scores, and autumn, winter, spring and summer growth scores generated a biplot (Figure 6.4). Directional vectors indicate the correlations among the traits in the biplot. The different seasons and years of the traits are shown in different colours, and the genotypes are indicated by red circles. PC1 and PC2 explained 85.94% and 6.86%, respectively, of the total variation (Figure 6.4). All traits were significantly associated with PC1 (Table 6.2). Stolon density and autumn, year 1, year 3, year 4 growth were significantly associated with PC2.

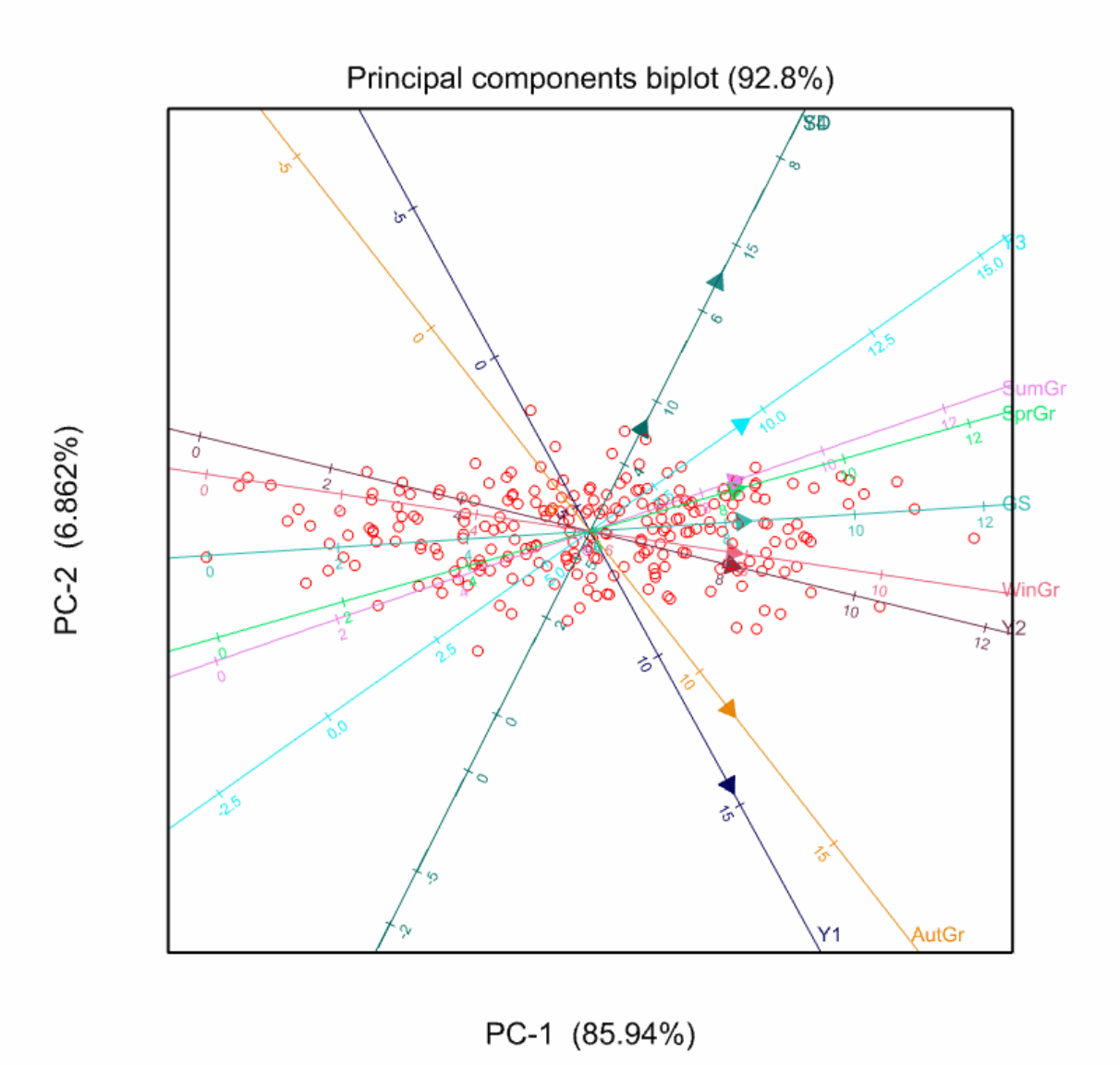

Figure 6.4 Biplot generated using standardised best linear unbiased predictor (BLUP)-adjusted means for stolon density (SD), overall growth score (GS), year 1 growth score (Y1), year 2 growth score (Y2), year 3 growth score (Y3), year 4 growth score (Y4), autumn growth score (AutGr), winter growth score (WinGr), spring growth score (SprGr) and summer growth score (SumGr) measured from 242 white clover genotypes. PC1 accounted for 85.94% and PC2 6.862% of the variation present. The directional vectors indicate the traits: Y1, Y2, Y3, Y4, AutGr, WinGr, SprGr, SumGr, SD and GS. Red circles indicate the genotypes.

Table 6.2 Correlation coefficients (*P*) and probabilities (*r*) for 10 plant traits with the first two principal components (PCs). The means of the traits were regressed against the scores of PC1 and PC2.

	Correlation with PC1		Correlation with PC2	
Trait description	*r*	*P*	*r*	*P*
Spring growth	0.975	0.001	0.079	0.223
Summer growth	0.972	0.001	0.095	0.139
Autumn growth	0.917	0.001	-0.332	0.001
Winter growth	0.971	0.001	-0.041	0.527
Year 1 growth	0.870	0.001	-0.449	0.001
Year 2 growth	0.931	0.001	-0.064	0.321
Year 3 growth	0.958	0.001	0.191	0.003
Year 4 growth	0.830	0.001	0.460	0.001
Growth average	0.999	0.001	0.018	0.781
Stolon density	0.611	0.001	0.339	0.001

PC analysis of the genotypes in the first subpopulation (Figure 6.9 cluster 1 and 2) showed that PC1 explained 85.84% of the total variation and PC2 explained 6.728% of the total variation (Figure 6.5). Meanwhile, PC analysis of the second subpopulation (Figure 6.9 cluster 3) showed that PC1 and PC2 accounted for a total of 93.43% of the total variation (Figure 6.6).

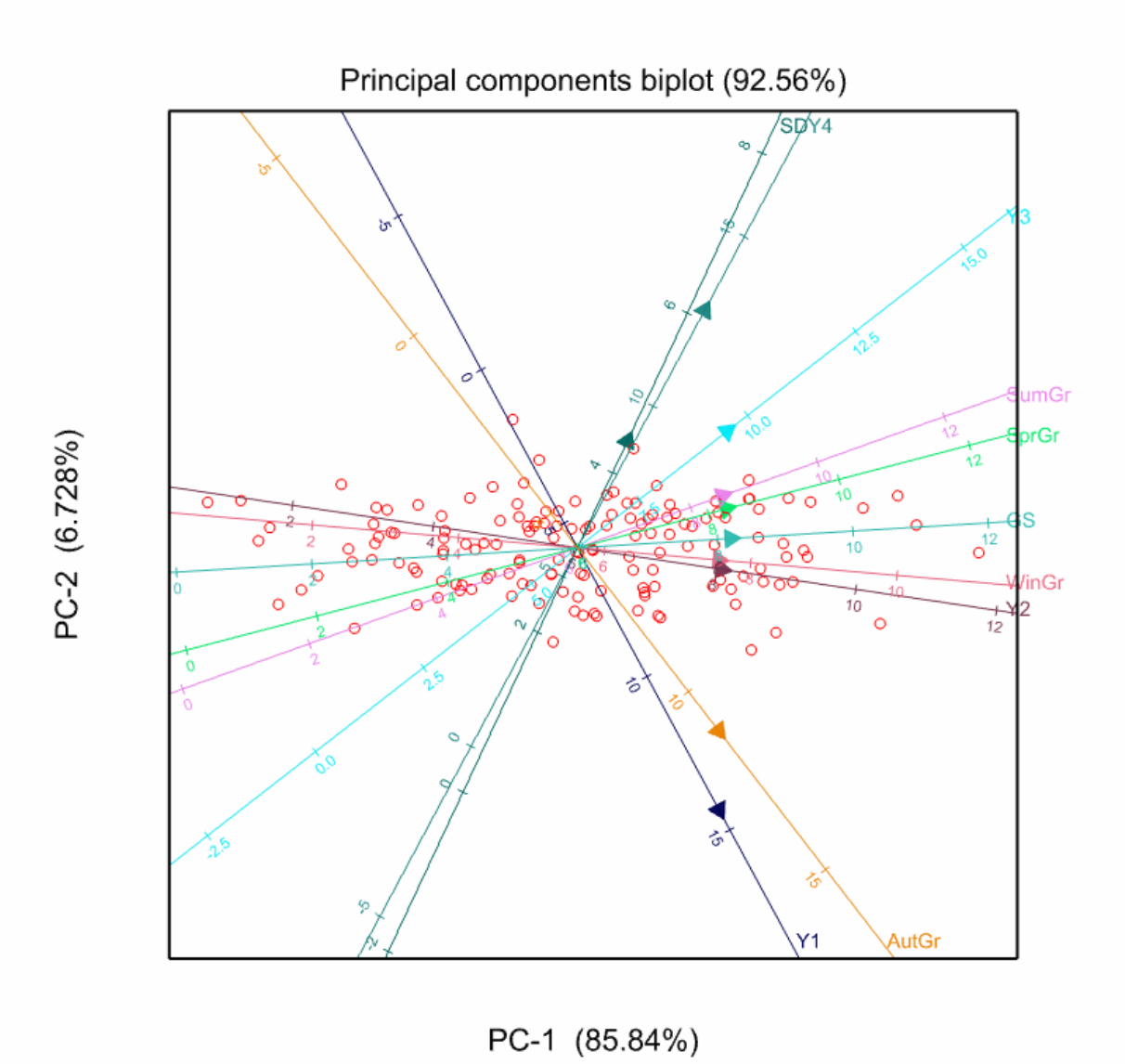

Figure 6.5 Biplot generated using standardised best linear unbiased predictor (BLUP)-adjusted means for stolon density (SD), overall growth score (GS), year 1 growth score (Y1), year 2 growth score (Y2), year 3 growth score (Y3), year 4 growth score (Y4), autumn growth score (AutGr), winter growth score (WinGr), spring growth score (SprGr) and summer growth score (SumGr) measured from the 150 white clover genotypes in the first subpopulation. PC1 accounted for 85.84% and PC2 6.728% of the variation present. The directional vectors indicate the traits: Y1, Y2, Y3, Y4, AutGr, WinGr, SprGr, SumGr, SD and GS. Red circles indicate the genotypes.

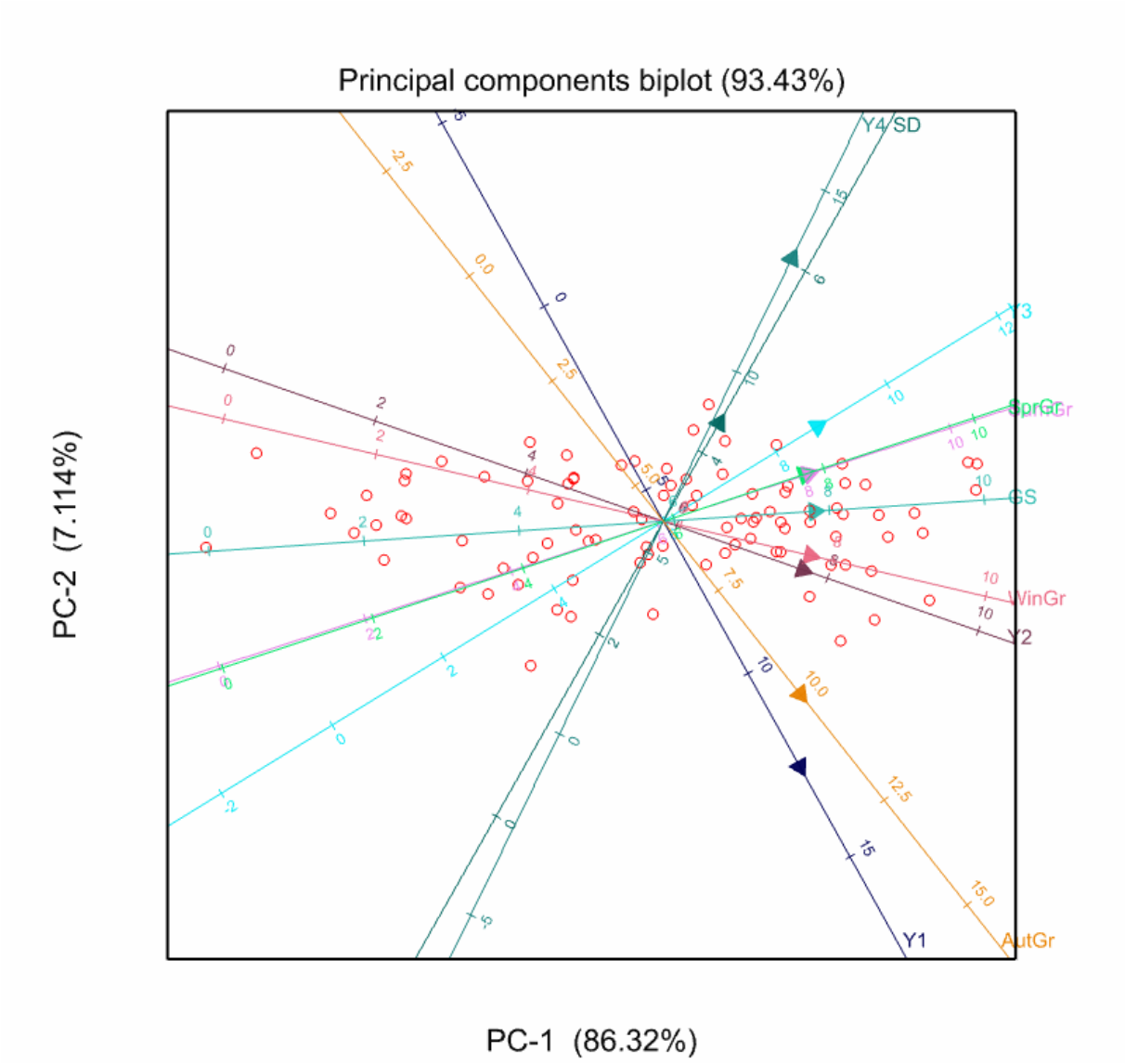

Figure 6.6 Biplot generated using standardised best linear unbiased predictor (BLUP)-adjusted means for stolon density (SD), overall growth score (GS), year 1 growth score (Y1), year 2 growth score (Y2), year 3 growth score (Y3), year 4 growth score (Y4), autumn growth score (AutGr), winter growth score (WinGr), spring growth score (SprGr) and summer growth score (SumGr) measured from the 92 white clover genotypes in the second subpopulation. PC1 accounted for 86.32% and PC2 7.114% of the variation present. The directional vectors indicate the traits: Y1, Y2, Y3, Y4, AutGr, WinGr, SprGr, SumGr, SD and GS. Red circles indicate the genotypes.

The estimates of correlations calculated for ten phenotypic traits are reported in Table 6.2 and visualised in Figure 6.7. All of the traits were correlated positively (p<0.001). In particular, strong correlations were found among growth score, autumn, winter, summer and spring growth. The correlations among seasons were, in general, stronger than among years. Years 2 and 3 growth showed strong correlations with spring, summer and winter growth. The weakest correlation (R=0.40) was between year 1 growth and stolon density. The correlation between the average growth score and stolon density was also weak (R=0.60).

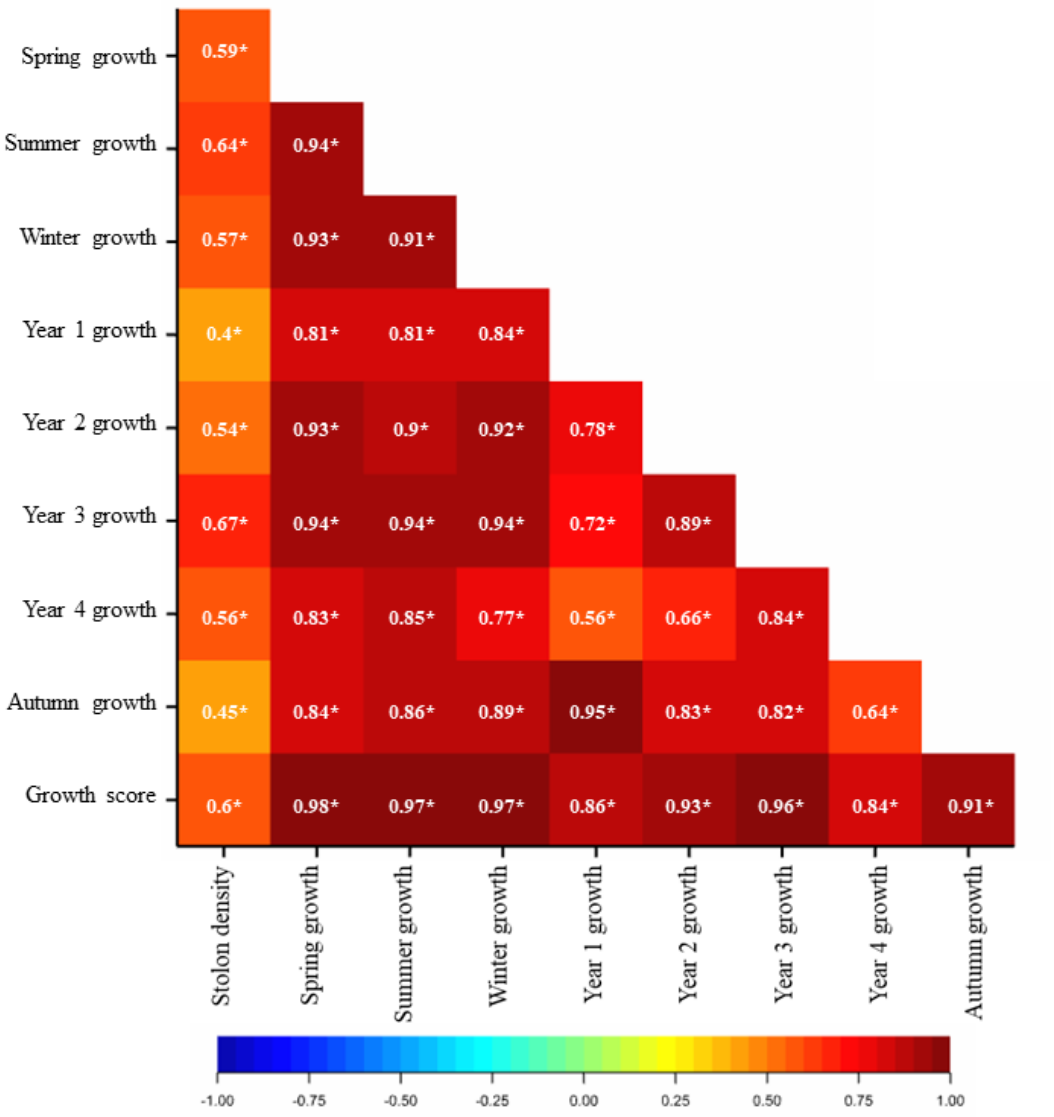

Figure 6.7 Correlation heatmap and correlation r values among the phenotypic traits of stolon density, spring growth, summer growth, winter growth, year 1 growth, year 2 growth, year 3 growth, year 4 growth, autumn growth and growth score for 242 white clover genotypes. The colours represent the correlation, with red being more positive and blue more negative. *P < 0·001.

6.3.3 Population structure and linkage disequilibrium

Population structure

Population structure effects are a frequent concern in GWAS studies. Accounting for population structure and relatedness is essential in mitigating false-positive associations (Sul, et al., 2018). fastStructure inferred population structure in the form of the optimal number of subpopulations (*k*) and resulted in a model complexity that had a maximum marginal likelihood of 2, and model components used to explain structure in the data of 2; or a *k* range of 1-2. ADMIXTURE used 10-fold cross-validation and calculated *k*=2. The most suited *k* value displayed the lowest cross-validation error (Figure 6.8). Hence, a *k* value of 2 was selected to describe the population structure.

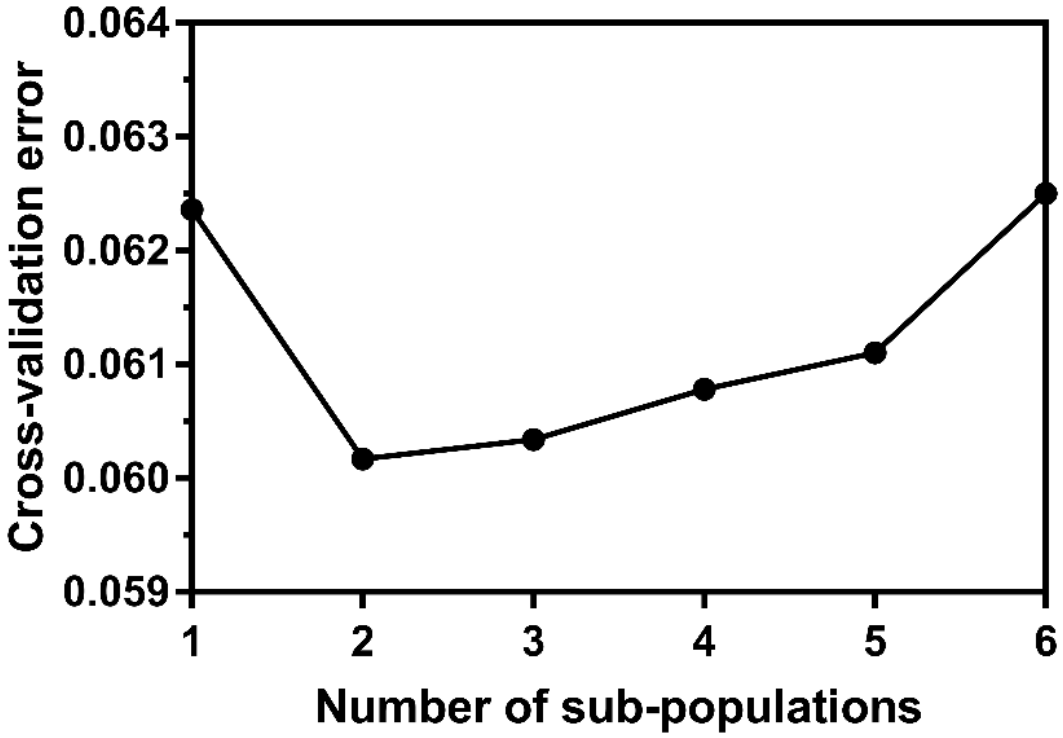

Figure 6.8 The number of subpopulations (1-6) simulated for the 242 white clover genotypes. The cross-validation error is on the y-axis, and the number of subpopulations is on the x-axis.

The neighbour-joining cladogram showed the genotypes clustering in three clusters (Figure 6.9). Cluster 1 had 139 genotypes, cluster 2 had 11 genotypes, and cluster 3 had 92 genotypes. The average growth and stolon density score in cluster 1 was 5.85 and 3.03, respectively (Table 6.3). Cluster 2 had the highest average stolon density score (3.42) and growth score (6.87). Meanwhile, the average growth score of cluster 3 was 5.86, and the average stolon density score was 3.25. The average growth and stolon density scores across all clusters were 5.90 and 3.13, respectively. However, the growth and stolon density scores were not significantly different between clusters (p>0.05).

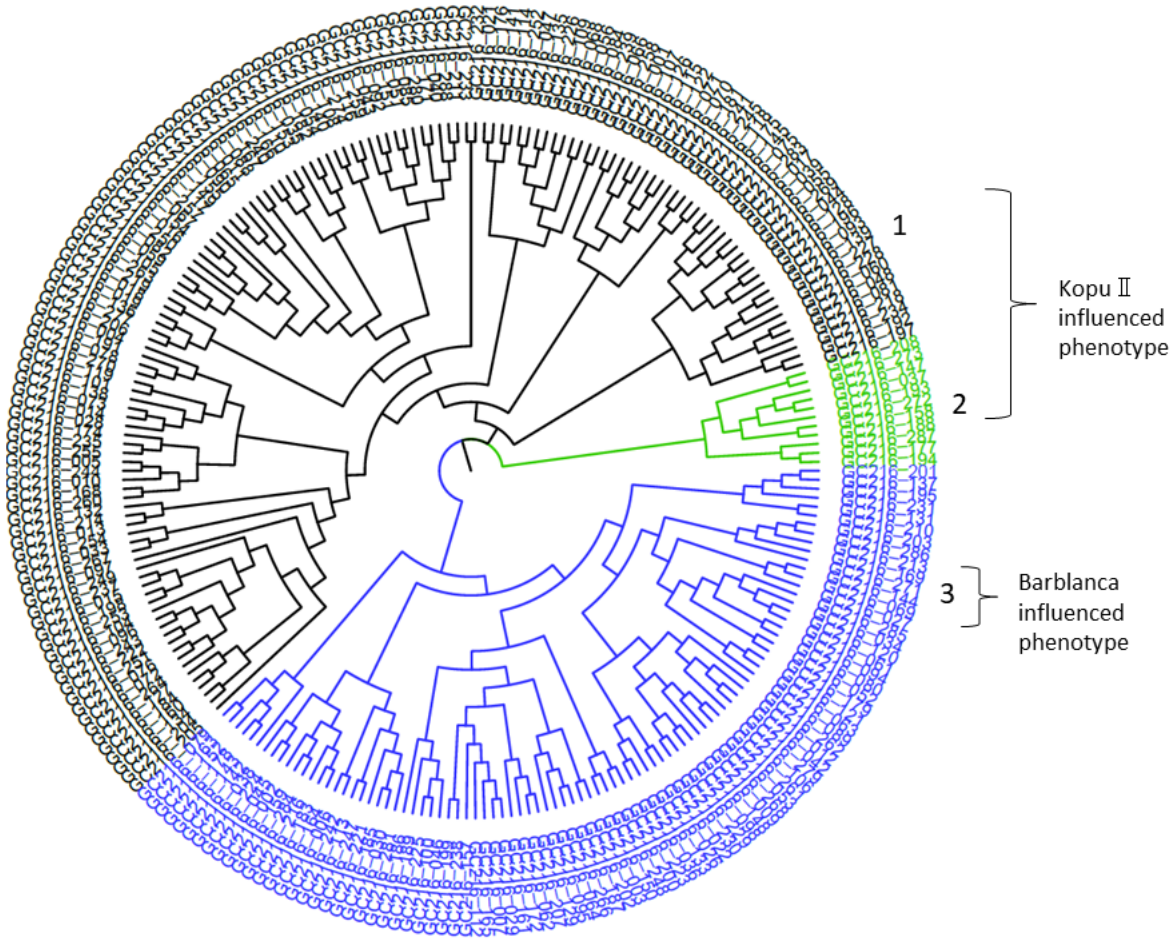

Figure 6.9 The neighbour-joining cladogram of the 242 white clover genotypes. The genotypes clustered into the three diverging groups; 1 (black), 2 (green) and 3 (blue). Clusters 1 and 2 were influenced by the Kopu II phenotype and cluster 3 was influenced by the Barblanca phenotype.

Table 6.3 The average growth and stolon density score and $LSD_{(0.05)}$ for the three clusters of the 242 white clover genotypes identified in the cladogram (Figure 6.9).

Cluster	Growth score	Stolon density score
$LSD_{(0.05)}$	1.102	0.616
1	5.85	3.03
2	6.87	3.42
3	5.86	3.25
Across all genotypes	5.90	3.13

Linkage disequilibrium

LD is the non-random association of alleles at different loci in a population. Many factors can alter LD including selection, population structure and genetic linkage. LD is measured by the comparisons of the observed frequencies of haplotypes to the frequencies expected based on the frequencies of the alleles that compromise the various haplotypes (Acquaah, 2012). For biallelic markers such as SNPs, LD commonly measured by r^2; the square of the correlation coefficient between two indicator

variables (Hill, et al., 1968). Genome-wide LD was estimated between pairs of SNPs across chromosomes. LD decay was rapid across the white clover panel as r^2=0.2 after 30bp.

6.3.4 Marker-trait associations

Associations were tested between 5,543 SNPs and each trait. Manhattan plots summarise the GWAS results in Figure 6.10. There were no significant SNPs for herbage yield and stolon density. The alignment of the QQ plots suggests that the pipeline did a sufficient job of controlling errors (Figure 6.11).

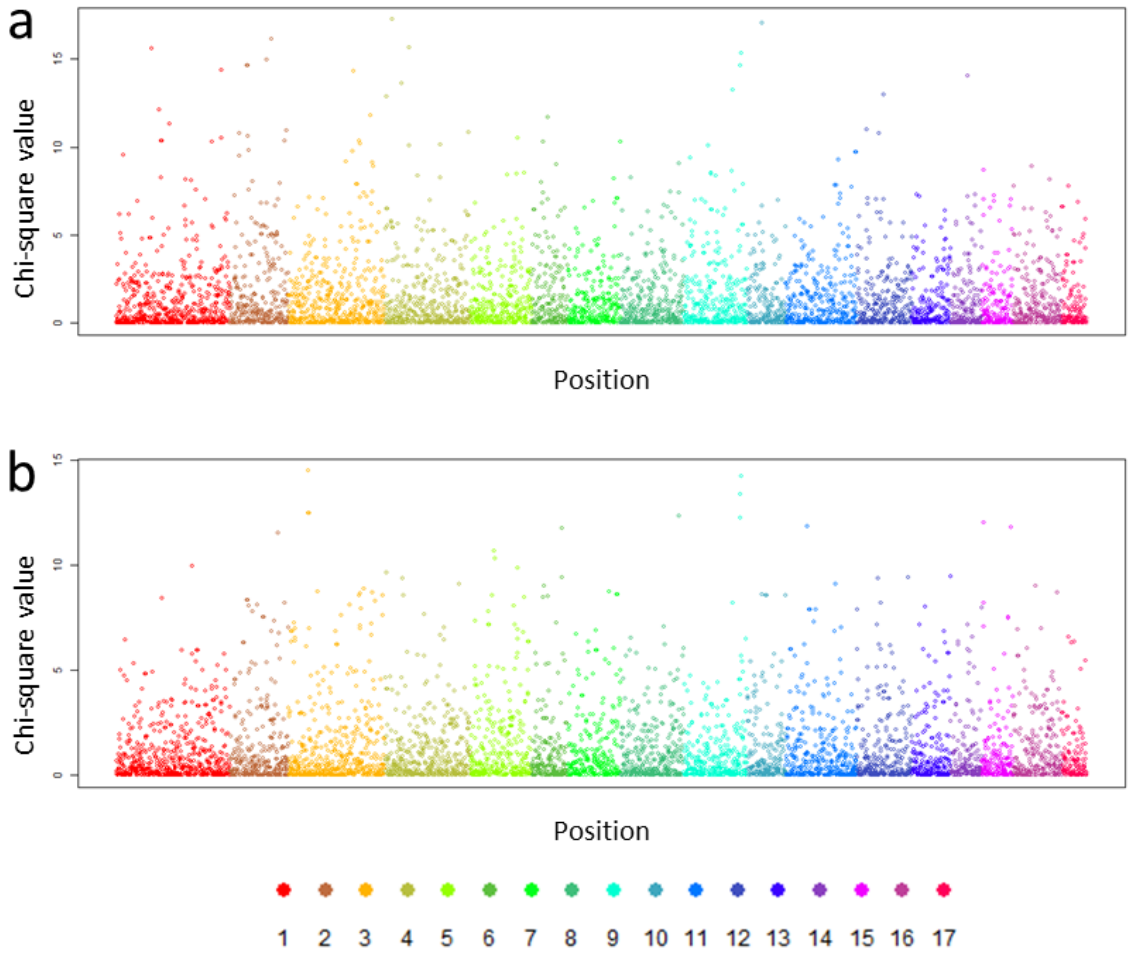

Figure 6.10 Manhattan plots for (a) stolon density, (b) herbage yield. The chi-square value is on the y-axis, and the position of the marker is on the x-axis. Colour indicates chromosomes 1-17.

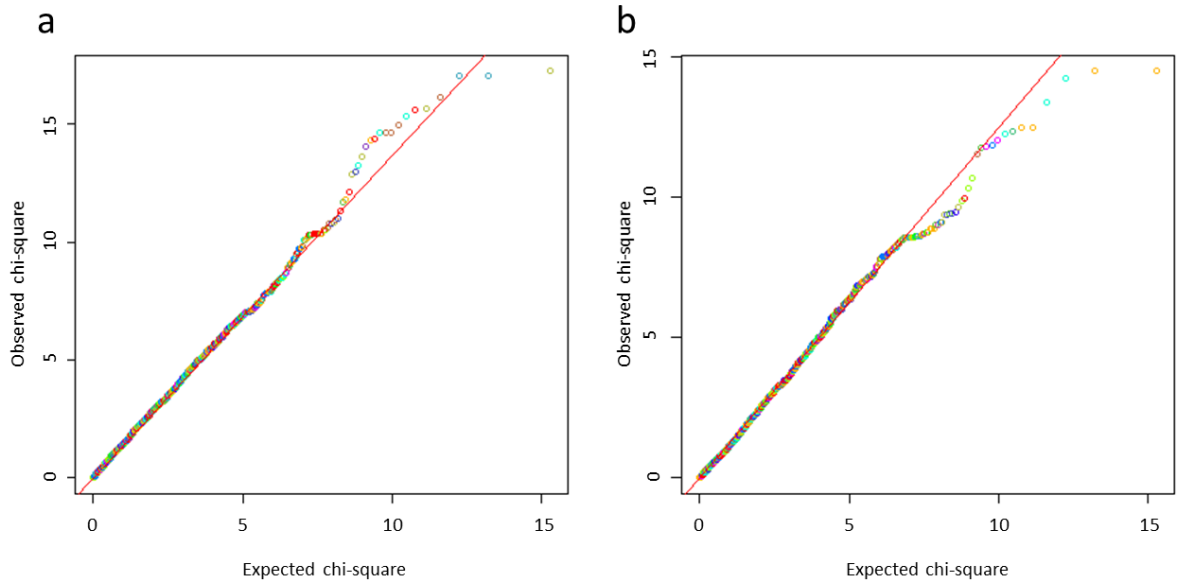

Figure 6.11 The quantile-quantile (QQ) plots for (a) stolon density, (b) herbage yield. The observed chi-square value is on the y-axis, and the expected chi-square value is on the x-axis.

6.4 Discussion

6.4.1 Population structure

Population structure is important to include in a mixed model to account for the relatedness among genotypes and to minimise Type 1 errors. In reality, a single *k* value from a population cannot be obtained as very rarely will a population be divided into discrete populations (Verity, et al., 2016). From a plant breeding perspective, it is important to identify the genetic structure and relatedness of the population. Population structure was estimated using two methods, and both calculated *k*=2. These results suggest that there is genetic differentiation among the genotypes. Population structure is common within populations in outcrossing plant species due to the highly heterozygous nature of the breeding systems. Several population genetics studies in white clover have presented similar results, and often within-population variation in white clover is present at high levels (Gustine, et al., 2001). Annicchiarico, et al. (2014) used simple sequence repeat markers on landrace and wild populations of white clover. While the variation between wild and landrace populations was low, the within-population variation was approximately 10-fold greater than among population variation.

The panel could be segregating into two subpopulations due to the breeding background of the population. The breeding material of the population was from a parental cross of the commercial cultivars Kopu II and Barblanca. The accessions associated with Kopu II and Barblanca in the MFGC were included in the pedigree analysis in Chapter 3 and were within the same cluster of the dendrogram (cluster 4, Figure 3.2). The pedigree traceback showed that Barblanca had a country of origin of The Netherlands and Kopu II had a country of origin of France. Kopu II had a large

percentage of the parentage from overseas germplasm including Italy, USA, Ireland and New Zealand germplasm (Woodfield, et al., 2001b). Whereas, Barblanca contained Spanish and Portuguese parentage. A study by Woodfield, et al. (2006) showed that Barblanca clustered with varieties of predominantly Mediterranean parentage. Both Barblanca and Kopu II had a strong influence of the Canterbury and Manawatu regions of New Zealand in the pedigrees. The pedigree analysis results in chapter 3 showed that the country of origin could strongly influence the phenotype of the lineage. As the parental germplasm of Kopu II is predominantly not from arid countries, whereas the Barblanca parental germplasm is, the difference in historical adaptation could be influencing the clustering pattern.

Phenotypic differences could be diverging the Mainstay panel. We could infer that the genotypes are clustering based on which genotypes belong more to a Grasslands™ Kopu II or Barblanca phenotype. The accessions in subpopulation 1 (clusters 1 and 2 in Figure 6.9) appear to have a Kopu II influenced phenotype. Subpopulation 1 outperformed subpopulation 2 in yield in the 3rd year but had the lowest stolon density. Widdup, et al. (2015) showed that Kopu II outperformed Barblanca in average yield under sheep grazing in the third year. The average growth score of subpopulation 1 was 6.36, which was higher than subpopulation 2 (5.86). Woodfield, et al. (1994) states that Kopu II was developed from overall high yielding populations (Mackay, 1991). Subpopulation 2 could be influenced by Barblanca due to the high stolon density (Table 6.3). Barblanca is marketed as having strong persistence compared to other large-leaved cultivars (Barenbrug, 2013).

Leaf size is a trait by which white clover cultivars are categorised. The leaf size of Barblanca is variable. Some reports state it is large-leaved (Woodfield, et al., 2006), while others state that it is medium/large-leaved (Widdup, et al., 2015). Kopu II has an undisputed large leaf size (Woodfield, et al., 2001b). In general, large-leaved cultivars tend to have lower persistence due to less stolon density (Mackay, 1991). The relatively open stolon network and long length of stolons allow for grazing damage and decreases persistence. Improved persistence is generally found in white clover cultivars that are small or medium leaf size because of the dense network of multi-branched stolons (Korte, et al., 1984). There is a well-known inverse relationship between leaf size and stolon density; the larger the leaf, the less dense the stolon network (Woodfield, et al., 1994). Overseas large-leaved germplasm has been bred with locally adapted small-leaved ecotypes to attempt to overcome the relationship. Although the relationship has not been decoupled, there has been an improvement in the persistence of medium-large leaved cultivars (Widdup, et al., 2011).

6.4.2 Linkage disequilibrium

LD determines the marker density required for marker-based studies. LD is a highly variable statistic and is dependent on the locus and the different selection pressures on different parts of the genome. Outcrossing species have a higher recombination rate which is a factor leading to fast LD decay. Higher levels of LD are observed in self-pollinating species compared to outcrossing species,

highlighting the role that the species mating system has on the level of LD. Tenaillon, et al. (2001) published the first study on LD in Maize before increased use of breeding methods that exploited inbreeding were used. LD decay was within 400bp. Comparatively, Hagenblad, et al. (2002) determined LD decay in a selfing *Arabidopsis* population to be ~250kb.

As white clover is a highly outcrossing species that suffers immensely from inbreeding depression, a rapid rate of LD decay is expected. Inostroza, et al. (2018) used a panel of 192 divergent genotypes of white clover from six populations. The LD showed rapid decay, with r^2 decreasing to 0.45 after 9kbp and below 0.2 after 134kbp. However, the LD decay may be overestimated as the *Medicago truncatula* (barrel clover) genome was used as the reference genome, which is smaller than the white clover genome. Many other reports of LD decay in outcrossing forage species report a rapid decay (Brazauskas, et al., 2010, Li, et al., 2014b). Sakiroglu, et al. (2012) investigated LD decay in 374 unimproved diploid alfalfa genotypes. In three of four candidate genes, LD decayed within 300bp below an r^2 of 0.2. Xing, et al. (2007) evaluated 20 genotypes of perennial ryegrass (*Lolium perenne* L.) and found that seven genes showed rapid LD decay of 400bp at r^2=0.2.

Hayes, et al. (2013) suggests that the rapid LD decay in outcrossing forage species could be due to a very large past effective population size. White clover is a recent allopolyploid, and Williams, et al. (2012b) first reported that the event arose approximately 13,000 to 130,000 years ago during the last glaciations. More recently, Griffiths, et al. (2019) estimated that white clover originated ~15,000 to 28,000 years ago. It is stated that Sved (1971) showed the expectation of r^2 is $\frac{1}{4N_e c+1}$ where N_e is the effective population size and c is the map distance (cM) between loci. Therefore, to have low r^2 values, a large N_e is required. White clover is bred through a variety of breeding techniques, such as biparental and polycrossing (Egan, et al., 2019a, Egan, et al., 2019b), which involve having large base populations. A common method of crossing is polycrossing, where many parents are placed in a crossing cage and pollinated with each other. This could infer the theory of a large N_e producing rapid LD decay. It is also a possibility that the mutation rate of white clover is so fast that the LD decay rate is rapid. Because effective recombination is increased in outcrossing species, LD will be less extensive. White clover is a species that is reliant on genetic drift and new variation in the population, so a rapid rate of LD decay is expected.

Although the SNPs were filtered for HWE, it is likely that the proportion of the SNPs in HWE equilibrium are not equal (Figure 6.1). The theoretical basis of estimating linkage disequilibrium is based on the assumption that the population is in HWE, and random mating occurs. When investigating at a genome-wide view, even though loci are individually in HWE, it may not be true when multiple loci are considered simultaneously (Liu, et al., 2012, Wu, et al., 2010). The assumption of random mating may be disrupted when exposed to evolutionary forces such as selection and population structure (Lynch, et al., 1998). The open-pollinated population in this study underwent as close as possible to random mating, yet still had an unequal proportion of SNPs in HWE. Populations

that have a small number of ancestors and are under constant selection can violate the assumptions of HWE. As this population was under constant selection and came from only two parental cultivars, and had a population structure of *k*=2, it could explain the uneven proportion of SNPs.

6.4.3 Marker-trait associations

The GWAS panel used in this study was a single breeding population rather than a diversity panel. Diversity panels are normally used in GWASs as the implementation of GWAS is crucially reliant on population diversity (Mackay, 2001). The ability for a GWAS to identify an association between a SNP and a trait is critically dependent on the phenotypic variance explained by the marker (Korte, et al., 2013). Increasing the panel size will have little effect on the ability to detect a signal unless the phenotypic variation is increased. The effect size, how strongly the two allelic variants differ in their phenotypic variance, and their frequency, determines the phenotypic variance, and many traits in outcrossing species are polygenic with small effect size (Korte, et al., 2013).

The phenotypic plasticity of white clover has been widely reported (Caradus, et al., 1993, Gautier, et al., 2001, Varin, et al., 2009, Welham, et al., 2002). The phenotypic plasticity is crucial for white clover production and survival under diverse climate environments and grazing management systems (Brock, et al., 1996). In this study, PC1 explained over 80% of the variation in all PCs, suggesting low levels of phenotypic diversity. The ability to detect a signal in a panel is significantly weakened when there are low levels of phenotypic variation. As previously explained, often in white clover populations, diversity is greater within a population than among populations. A recent report by Gage, et al. (2017) suggests that the improvement of maize by breeding may have reduced the genetic control of phenotypic plasticity through genotype by environment (GxE) interaction. This phenomenon could be contributing to the inability to find signals in a large white clover panel.

All of the traits assessed in this study were found to be significantly associated with PC1. However, only stolon density and autumn, year 1, year 3, year 4 growth were significantly associated with PC2. White clover cultivars are bred to perform across many environments and for many years. During the breeding programmes, the cultivars are assessed in grazing trials for a minimum of three years, with a total of ten years of evaluation before they are released to the market. The relationship of the traits with PC1 suggests that the population has consistently performed across the four years of evaluation. The weakest correlation among the traits was between year 1 growth and stolon density. In the first year of growth for white clover, there is a low stolon density as the taproot is still present. After ~18 months, the taproot dies, and the stolons are solely reliant on the nodal roots (Widdup, et al., 2011). The weak correlation between year 1 growth and stolon density is explained by the young growth stage of the white clover plants.

The average growth score and stolon density score had a correlation of 0.60. Although survival of white clover is dependent on the persistence provided through a strong stolon density network, it is

well documented in the literature of the negative and often weak association between stolon density and herbage yield (Caradus, et al., 1989c, Jahufer, et al., 1999). As the growth score is used to infer herbage yield, this correlation supports the literature. However, the literature is unclear as to the basis of the negative correlation. Possible reasoning is pleiotropy or linkage, which could be overcome by breeding techniques, i.e. selecting recombinant genotypes (Jahufer, et al., 1999).

It is well documented in the major cereals of wheat, barley and rice, that yield-related traits are controlled by large, complex gene networks (Gasparis, et al., 2011, Hanif, et al., 2016, Houston, et al., 2013, Huang, et al., 2018, Nadolska-Orczyk, et al., 2017, Slavov, et al., 2014, Yu, 2017, Zheng, et al., 2014). However, although the stolon characteristics of forages have been poorly uncharacterised in a molecular means (Cnops, et al., 2010, Inostroza, et al., 2018), it is thought that they are complex. Complex traits are hard to dissect as rare variants, and small effects sizes are often not detected in GWASs (Asimit, et al., 2010, Gibson, 2012). Dissecting complex traits through GWAS requires both high levels of phenotypic variation and power. In an attempt to mitigate the difficulties of dissecting complex traits, often GWASs are performed on species exposed to extreme climates (Ma, et al., 2016, Pham, et al., 2019, Tessmann, et al., 2018, Volante, et al., 2017).

6.5 Conclusions

- PC analysis of the ten traits identified that all traits had a significant association with PC1, while some had a significant association with PC2. The correlation of the traits showed strong correlations within both years and seasons but highlighted the weak correlation between stolon density and growth score.

- The results from this chapter show the challenges of determining the genetic architecture of complex traits such as herbage yield and stolon density.

- The lack of significant SNPs is most likely attributed to the lack of phenotypic variation within the population. The levels of phenotypic variation would need to be increased to increase the likelihood of detecting a signal.

- The population structure of k=2 is most likely segregating because of the two parental cultivars. The high levels of genetic variation seen within the population are similar to other studies of within-population variation in white clover.

- The rapid rate of LD decay was expected due to the outcrossing and highly heterozygous nature of white clover.

Chapter 7
Transpiration rate of white clover (*Trifolium repens*) cultivars in drying soil

This chapter has been prepared for submission to Frontiers in Genetics, special edition High-Throughput Phenotyping for Crop Improvement and Breeding.

7.1 Introduction

White clover is the most important pastoral legume in temperate regions of the world and is usually grown in companion with ryegrass. The pasture mix of ryegrass and white clover is common in a variety of grazing systems, including sheep and beef, deer, and dairy. Globally, white clover is an attractive plant to have in pastoral systems due to the nitrogen fixation ability and the resulting role in sustainable farming systems. White clover is economically important to New Zealand and fixes approximately 1.57 million tonnes of nitrogen annually. New Zealand has the highest export share of white clover seed globally (57.5%), exporting approximately 4500 tonnes of white clover seed annually (Rattray, 2005).

Globally, drought stress is one of the major limiting factors in white clover performance in pastoral systems. Climate change predictions show that there will be a global increase in temperature between 4.5˚C to 18˚C over the next century; increasing drought periods globally (Fischlin, et al., 2007, IPCC, 2013). With the expansion of farming into arid geographic areas, climate change and increased water restrictions, there has been more urgency to breed and utilize cultivars that can perform under drought stress.

There have been several studies investigating the response of white clover to drought conditions (Annicchiarico, et al., 2004, Bermejo, et al., 2006, Li, et al., 2013a, Thomas, 1984). Thomas (1984) showed that the competitiveness of white clover in a sward is reduced significantly in drought-stressed environments which impacts severely on-farm production. There has been some success in studies investigating the biochemical and molecular basis of drought effects in white clover (Li, et al., 2014c, Li, et al., 2013b). Genomically, drought tolerance is extremely complex and a difficult trait to understand (Azhar, et al., 2018, Blum, 2011). Drought tolerance as a trait is polygenic and has been shown across many species to involve a large number of genes and different gene expression between plant tissues and growth stage (Blum, 2011, Chen, et al., 2008, Fleury, et al., 2010, Jiang, et al., 2010, Kasuga, et al., 2004, Shinozaki, et al., 2007, Umezawa, et al., 2006, Xiong, et al., 2006).

Recurring droughts in New Zealand have increased farm management challenges for farmers. There have been many studies aimed at increasing performance in drought through agronomic practices (van den Bosch, et al., 1993, Widdup, et al., 2011). Although there has been little success in breeding

drought tolerant white clover cultivars, differing grazing management schemes can aid in protecting the plants under drought conditions (Brock, 1988).

Understanding the relationship between drought tolerance and persistence will be beneficial for the further development of cultivars (Sanderson, et al., 2003). Persistence is defined as the maintenance of long-term agronomic yield and is a function of stolon growth and density (Annicchiarico, et al., 2004, Williams, 1987). Drought is one of the main persistence-limiting traits (Bouton, 2012). Drought conditions limit white clover stolon branching and rooting (Chapman, 1983), and higher stolon density is often associated with superior water conservation (Collins, 1998). The success of a forage cultivar is largely dependent on the ability of the plant to survive summer droughts and retain productivity (Annicchiarico, et al., 2011, Pecetti, et al., 2011). Both persistence and stolon density can be traits of interest to increase drought tolerance. Hutchinson, et al. (1995) carried out a 30-year study of environmental factors affecting the persistence of white clover and found that drought stress in late summer was the most critical limiting factor. Belaygue, et al. (1996) saw an 80% decrease in stolon density when rainfall decreased by 30%. Saeidnia, et al. (2018) noted that drought-tolerant genotypes of orchardgrass had high persistence and Saeidnia, et al. (2016) showed that drought conditions reduced forage yield and persistence strongly in orchardgrass.

Traditional breeding methods, such as phenotypic selection, have been the most common in white clover breeding programmes. Selecting for deeper and more extensive rooting systems is a common breeding strategy. However, the correlation between root depth and drought tolerance remains unclear. The results from studies which analysed the association of rooting depth and drought tolerance are mixed. Caradus (1981) reported that germplasm with larger leaf size and root systems outperformed germplasm with smaller root systems. However, Barbour, et al. (1996) found no significant difference between the market class of the cultivar and the performance under drought conditions. Selecting for a deeper root system is complicated by the rooting pattern of white clover. A taproot is present for the early life of the plant (12-18 months) and afterwards is replaced with a shallow nodal root system. Annicchiarico, et al. (2004) proposed an alternative in selecting for greater root systems by selecting for thicker stolons and found that the plants with thicker stolons had increased root dry weight.

White clover cultivars need to be exposed to the same drought condition to make accurate assumptions on differences of phenotypic performance in drought. In a field trial environment, phenotypic differences in performance can be due to differences in water regimes. Other measures of drought tolerance are needed to accurately determine the performance of plants in drought conditions (Ray, et al., 1997, Ray, et al., 1998, Sinclair, et al., 1986, Weisz, et al., 1994).

Selecting genotypes that are conservative in water use is a selection strategy to increase drought tolerance. Traits such as normalised transpiration rate (NTR) and the fraction of transpirable soil water (FTSW) can be used to determine cultivars or families that perform better in drought-like conditions (Lecoeur, et al., 1996, Miller, 2000, Ray, et al., 1997, Ray, et al., 1998). Genotypes that are more

sensitive to drought will close their stomata earlier to preserve soil water content and may perform better in long-period drought conditions as water-conserving efforts occur earlier. Genotypes that have late stomatal closure may be better suited for quick-period drought conditions.

Currently, there is limited published literature on calculating the NTR and FTSW critical threshold (FTSWc) of white clover cultivars in drying soil. We used a panel of 80 white clover cultivars to determine the FTSWc, marking permanent stomatal closure and the start of senescence.

7.2 Materials and Methods

7.2.1 Germplasm

The 80 cultivars used in this study were the same as in Hoyos-Villegas, et al. (2019) and were released from 1920-2010 by both public and private breeding programs. The cultivars were from across New Zealand, Australia, United Kingdom and the United States of America. The cultivars ranged in leaf size of small (N=1), medium (N=53) or large (N=26) (Table 7.3).

7.2.2 NTR and FTSW trial design

Two glasshouse experiments were conducted at AgResearch, Lincoln. The first experiment was run from 09/12/2016-22/12/2016. Weather and glasshouse sensor data for that period showed that the average temperature was 22.39˚C, the average relative humidity was 57.00%, the average total daily solar radiation was 24.39MJ/m^2, and the average vapour pressure deficit was 0.39kPa (Table 7.1). The second experiment was run from 02/02/2017-20/02/2017. The average temperature was 22.79˚C, the average relative humidity was 56.51%, the average total daily solar radiation was 19.82MJ/m^2, and the average vapour pressure deficit was 0.41kPa (Table 7.2). The weather station was not equipped to measure sunshine hours.

Table 7.1 The relative humidity (%), mean temperature (˚C), daily solar radiation (MJ/m^2) and vapour pressure deficit (kPa) from a weather station and glasshouse sensors for the first experiment (09/12/2016 - 22/12/2016).

	Relative humidity (%)	Mean temperature (˚C)	Daily solar radiation (MJ/m^2)	Vapour pressure deficit (kPa)
9/12/2016	64.65	23.02	23.98	0.33
10/12/2016	58.82	23.35	20.93	0.39
11/12/2016	63.71	20.58	19.81	0.31
12/12/2016	59.48	20.35	22.14	0.34
13/12/2016	54.41	24.56	24.70	0.45
14/12/2016	52.85	23.40	20.84	0.45
15/12/2016	52.79	23.16	29.84	0.44
16/12/2016	53.11	21.91	30.57	0.42
17/12/2016	60.97	21.77	32.02	0.35
18/12/2016	65.15	20.95	9.27	0.30
19/12/2016	50.55	22.58	33.40	0.45
20/12/2016	55.55	20.59	31.23	0.37
21/12/2016	48.22	25.50	32.26	0.53
22/12/2016	57.80	21.72	10.44	0.37
Average	**57.00**	**22.39**	**24.39**	**0.39**

Table 7.2 The relative humidity (%), mean temperature (˚C), daily solar radiation (MJ/m^2) and vapour pressure deficit (kPa) from a weather station and glasshouse sensors for the second experiment (02/02/2017 - 20/02/2017).

	Relative humidity (%)	Mean temperature (˚C)	Daily solar radiation (MJ/m^2)	Vapour pressure deficit (kPa)
2/02/2017	62.61	20.52	11.46	0.31
3/02/2017	53.60	23.98	25.61	0.45
4/02/2017	51.54	26.30	21.87	0.51
5/02/2017	52.93	26.29	26.15	0.50
6/02/2017	54.93	25.22	13.23	0.46
7/02/2017	64.33	19.50	5.24	0.29
8/02/2017	54.37	21.59	20.77	0.40
9/02/2017	48.45	21.25	27.74	0.45
10/02/2017	55.06	21.56	17.19	0.39
11/02/2017	52.95	24.19	26.08	0.46
12/02/2017	53.91	24.38	16.28	0.45
13/02/2017	51.24	23.31	20.32	0.46
14/02/2017	34.60	24.48	27.92	0.65
15/02/2017	39.57	22.46	27.03	0.55
16/02/2017	52.46	23.66	25.43	0.45
17/02/2017	69.96	20.53	7.66	0.25
18/02/2017	68.28	21.34	11.16	0.28
19/02/2017	67.42	24.11	20.96	0.32
20/02/2017	85.54	18.30	24.55	0.11
Average	**56.51**	**22.79**	**19.82**	**0.41**

The 80 cultivars were exposed to two treatments, irrigated (control) and drought, and replicated twice in a randomised complete block design. Each cultivar was potted individually and grown in a glasshouse with the optimum temperature of 20-25˚C maintained. Ten bare pots were used to measure water loss through evaporation. The pots were watered to field capacity and then left to drain and sealed. The saturated weight of each pot was recorded. Daily measurements of soil water content, the weight of each pot, and the amount of water transpired, measured as the difference in pot weight, were recorded.

Dry weight was measured by harvesting the roots and shoots. The samples were dried in the oven at 80˚C for 12 hours and weighed (g). The leaf size used was the commercially stated leaf size associated with the cultivar.

7.2.3 Calculations

Vapour pressure deficit (VPD) was calculated using the formula in Conaty, et al. (2014):

$$VPD = e_s - e_a$$

where VPD is the difference between ambient water vapour (e_a) and the saturated vapour pressure (e_s) at the same temperature. The air temperature (T_a in ˚C) and relative humidity (RH in %) where:

$$e_s = 0.6108e\left(\frac{17.27T_a}{T_a + 237.3}\right)$$

$$e_a = \left(\frac{RH}{100}\right)e_s$$

The transpiration data were analysed by the methodology described in Ray, et al. (1997).

The transpiration rate (TR) was calculated using the formula:

$$TR = \frac{Weight\ of\ drought\ pot}{Weight\ of\ control\ pot}$$

where the soil water content and amount of water transpired measurements were normalised against the control pots.

The NTR was calculated using the formula:

$$NTR = \frac{TR}{Day\ 3\ and\ 5\ average\ TR}$$

The transpiration values were normalised against the day 3 and 5 average TR to minimize the effects of fluctuations in transpiration. The TR on days 3 and 5 are considered to be under well-watered

conditions, allowing the plants to have an average NTR near-equal to one when sufficient soil water was available.

FTSW was calculated using the formula:

$$Daily\ FTSW = \frac{(Daily\ pot\ weight - Final\ pot\ weight)}{(Initial\ pot\ weight - Final\ pot\ weight)}$$

The relationship between the transpiration value and FTSW of each cultivar was explained by using non-linear regression to fit the equation:

$$NTR = \frac{1}{[1 + A \times exp(B \times FTSW)]}$$

Comparisons of the curve generated for each cultivar were based on 95% confidence intervals of coefficients *A* and *B*. Plateau regression was used to determine the FTSWc. The curve predicts that NTR will remain near 1 up until a critical point, after which NTR decreases. The FTSWc is estimated, where NTR decreases linearly after that. The FTSWc marks the critical point where the stomata start to close, and transpiration decreases linearly.

7.2.4 Statistical analysis

Analysis of variance (ANOVA) in GenStat (VSN-International, 2019) examined the main effects of water and cultivar, as well as their interaction. Before ANOVA, data were tested for homogeneity of variances and no transformation of the data was required. Fishers LSD at $p<0.05$ was used throughout the chapter.

A linear spline model was used to analyse the trend of the FTSWc over decades. Since the average trend showed a linear increase from 1920 up to 1960, then, changed into a decrease after 1960, the trend pattern was modelled as a linear spline model with a single knot at 1965. The linear spline trend model was estimated in regression by the formula:

$$Inflection\ point = a + b\ x\ Time + c\ x\ Time\_2$$

where *Time* is a variable representing the decades in order, *Time_2* is a variable also representing decades in a different order, and *a*, *b* and *c* are parameters to be estimated. In this formula, the increasing trend up to 1960 was estimated as a positive *b* value - i.e. *b* = increasing rate per decade till 1960, while the decreasing trend after 1960 was estimated as the sum of *b* + *c*, which should be a negative value – i.e. the absolute value of *b* + *c* = decreasing rate per decade after 1960. In addition, since the value of *c* indicated the difference of two trend slopes (one till 1960 and the other after 1960), the significance level (*p*-value) associated with *c* indicated if the change of the trend from the increase to the decrease was statistically significant.

7.3 Results

7.3.1 Dry weights

The average dry weight for drought and irrigated plants was 16.27g and 22.70g, respectively. The plant total dry matter weights ranged from 11.23g to 29.25g for control plants and 10.65g and 23.32 for drought plants (Table 7.3 and Figure 7.1). There was a significant overall difference ($p<0.01$) in the dry weights of the cultivars when averaged across the two water treatments (Duncan, 1955). The medium-leaved cultivars released in the 1920s were significantly different from all other groups for drought and irrigated plant dry weight (Tables 7.4 and 7.5). The large-leaved cultivars released in the 1950s and 1960s were significantly different from all other groups for drought plant dry weight (Table 7.4). The cultivar and drought interaction was not significant.

Table 7.3 The cultivar names, the decade of release, countries of origin, registered leaf sizes, dry weights of the drought and irrigated treatments and the critical fraction of transpirable soil water (FTSWc) threshold of eighty white clover cultivars as indicated in Hoyos-Villegas, et al. (2019). Fishers $LSD_{0.05}$ is presented for the dry weights of the irrigated and drought cultivars and the FTSWc. The '/' denotes that the FTSW threshold was unable to be calculated for the cultivar because of the irregularities in the data for curve generation. Abbreviations: NZ, New Zealand; UK, United Kingdom; USA, United States of America.

Cultivar	Decade of release	Country of origin	Leaf size	Dry weight drought	Dry weight irrigated	FTSWc
$LSD_{(0.05)}$				1.99	1.99	2.02
Dutch white	1920	Netherlands	Medium	13.58	17.75	0.23
Irrigation	1930	Australia	Medium	18.63	20.78	/
Kent White	1930	UK	Small	16.85	26.28	0.17
Louisiana	1930	USA	Medium	17.78	29.08	0.25
S 100	1930	UK	Medium	18.73	20.70	0.35
Kersey	1940	UK	Medium	21.55	28.68	/
S 184	1940	UK	Medium	17.18	23.05	0.28
California Ladino	1950	USA	Large	15.65	23.18	/
Grasslands Huia	1950	NZ	Medium	17.45	25.95	0.36
Ladino Gitante Lodigiano	1950	USA	Large	23.33	27.40	/
Louisiana S1	1950	USA	Medium	20.53	26.25	/
Pilgrim	1950	USA	Medium	17.35	29.25	0.28
Sonja	1950	Sweden	Large	18.70	23.30	0.33
Tribla	1950	Belgium	Medium	13.70	21.38	0.42
Clarence	1960	Australia	Medium	15.88	19.83	0.22
Crau	1960	France	Medium	16.43	22.38	/
Haifa	1960	Israel	Medium	14.95	22.35	0.38
Regal	1960	USA	Large	13.80	20.23	/
Donna	1970	UK	Medium	13.75	19.78	0.37
Lune de mai	1970	France	Large	17.28	24.23	0.21
Milkanova	1970	Denmark	Medium	17.18	24.13	0.26
Olwen	1970	UK	Large	15.08	21.88	/

Pitau	1970	NZ	Medium	14.00	19.20	0.48
Radi	1970	Poland	Large	15.85	24.23	0.23
Sacramento	1970	Poland	Large	15.85	21.48	0.49
Siral	1970	Australia	Medium	15.25	18.35	/
Alice	1980	UK	Medium	14.18	19.38	0.36
Aran	1980	Ireland	Large	13.30	22.78	/
Kopu	1980	NZ	Large	17.05	21.93	/
Lirepa	1980	Germany	Medium	17.65	22.83	0.34
Menna	1980	UK	Medium	12.23	21.90	0.39
Merwi	1980	Belgium	Medium	15.35	22.33	0.37
Osceola	1980	USA	Medium	12.93	17.20	0.42
Ross	1980	Ireland	Large	16.75	24.95	0.40
AberHerald	1990	UK	Medium	14.13	22.25	0.12
Challenge	1990	NZ	Medium	14.40	21.40	0.27
Crescendo Ladino	1990	USA	Large	12.88	19.75	0.32
Dacia	1990	Romania	Large	16.53	26.03	0.28
Jumbo	1990	USA	Medium	18.80	24.00	0.20
Kopu II	1990	NZ	Large	14.33	20.80	0.34
Le Bons	1990	NZ	Medium	16.10	18.80	0.24
Prop	1990	NZ	Medium	14.20	23.00	0.34
Regal Graze	1990	USA	Large	18.15	28.40	/
Riesling	1990	Netherlands	Medium	15.75	22.75	0.30
Sustain	1990	NZ	Medium	16.05	17.50	0.32
Triffid	1990	France	Large	16.23	24.30	0.30
Waverley	1990	Australia	Large	12.93	22.53	/
AberConcord	2000	UK	Medium	15.95	23.13	0.20
AberDance	2000	UK	Medium	17.88	22.65	0.35
AberNormous	2000	UK	Large	16.78	21.58	0.26
Aquiles	2000	Uruguay	Medium	20.50	24.40	0.13
Artigas	2000	Uruguay	Large	16.10	22.90	/
Barblanca	2000	France	Medium	15.75	25.08	0.27
Bounty	2000	NZ	Medium	12.53	20.43	0.47
Chieftain	2000	Ireland	Medium	17.25	21.40	0.17
Crusader	2000	France	Medium	15.60	16.40	/
Emerald	2000	NZ	Medium	15.98	20.78	0.33
Goliath	2000	Uruguay	Large	17.95	26.58	0.17
Klondike	2000	Denmark	Medium	13.95	25.33	/
Kotare	2000	NZ	Large	14.65	19.45	0.42
Quest	2000	NZ	Medium	13.95	20.95	0.33
Saracen	2000	Australia	Medium	15.28	24.20	0.27
Super Haifa	2000	Australia	Medium	15.93	21.75	0.24
Super Ladino	2000	Australia	Large	17.10	22.98	0.18
Tasman	2000	NE	Medium	13.48	21.80	0.30
Tillman II	2000	USA	Large	16.80	23.15	0.28
Tribute	2000	NZ	Medium	16.48	21.05	0.52
Trophy	2000	Australia	Medium	10.65	11.23	/

Vysocan	2000	Czech	Large	15.33	21.93	/
ABM21252	2010	NZ	Large	17.33	24.25	0.18
Calimero	2010	USA	Medium	17.93	22.50	/
Dairy B GC276	2010	Australia	Medium	16.18	21.58	/
Dairy D	2010	NZ	Medium	21.53	26.08	0.18
Elite Breeding A	2010	Australia	Medium	18.05	24.73	0.24
Kakariki	2010	NZ	Large	18.08	25.93	/
Katy	2010	USA	Medium	18.70	21.28	0.24
Legacy	2010	NZ	Large	18.15	27.98	0.29
Mainstay	2010	NZ	Medium	18.40	26.95	0.29
Quartz	2010	NZ	Medium	16.53	23.58	0.24
Weka	2010	NZ	Medium	18.75	25.98	0.18

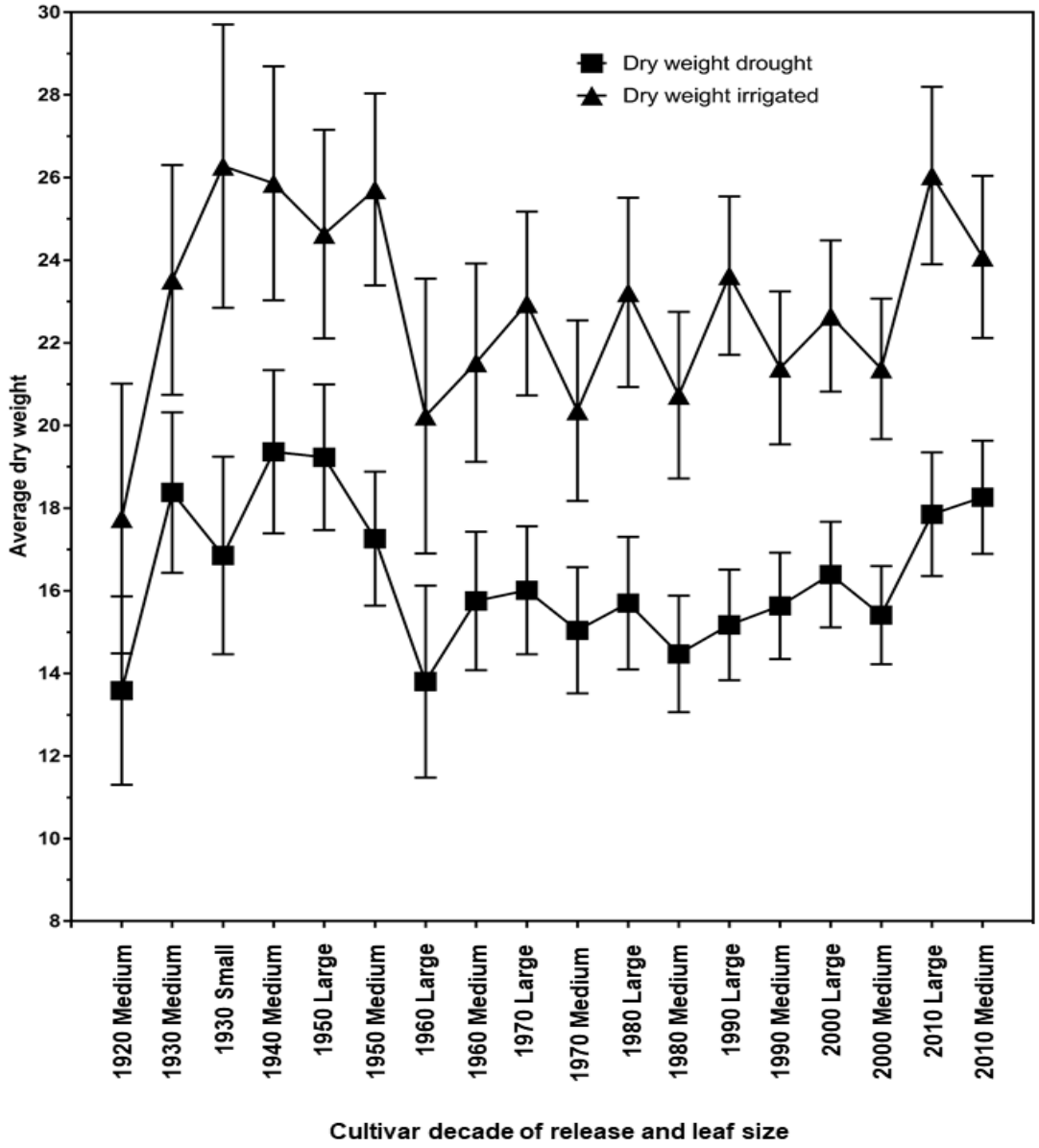

Figure 7.1 Average dry weights for irrigated and drought treated plants, clustered by leaf size and decade of release.

Table 7.4 The average dry weight for drought exposed cultivars, clustered by leaf size and decade of release. Duncan's lettering compared the dry weights for leaf size and decade of release and the groups without a common small letter are significantly different at the 5% level of probability.

Decade of release and leaf size	Dry weight average	$LSD_{0.05}$
1920 Medium	13.58a	4.57
1960 Large	13.8a	4.65
1980 Medium	14.47ab	2.82
1970 Medium	15.04abc	3.05
1990 Large	15.17abc	2.68
2000 Medium	15.41abc	2.38
1990 Medium	15.63abc	2.58
1980 Large	15.7abcd	3.21
1960 Medium	15.75abcde	3.36
1970 Large	16.01abcde	3.10
2000 Large	16.39abcde	2.56
1930 Small	16.85abcdef	4.79
1950 Medium	17.26acdef	3.25
2010 Large	17.85acdef	3.00
2010 Medium	18.26def	2.75
1930 Medium	18.38def	3.89
1950 Large	19.23f	3.53
1940 Medium	19.36ef	3.96

Table 7.5 The average dry weight for irrigated cultivars, clustered by leaf size and decade of release. Duncan's lettering compared the dry weights for leaf size and decade of release and the groups without a common small letter are significantly different at the 5% level of probability.

Decade of release and leaf size	Dry weight average	$LSD_{0.05}$
1920 Medium	17.75a	6.54
1960 Large	20.23ab	6.66
1970 Medium	20.36abc	4.37
1980 Medium	20.73abc	4.03
2000 Medium	21.37abc	3.41
1990 Medium	21.39abcd	3.70
1960 Medium	21.52abcde	4.81
2000 Large	22.65abcde	3.67
1970 Large	22.95abcde	4.44
1980 Large	23.22abcde	4.59
1930 Medium	23.52abcde	5.56
1990 Large	23.63abcde	3.84
2010 Medium	24.08bde	3.93
1950 Large	24.63bcde	5.05
1950 Medium	25.71be	4.65
1940 Medium	25.86bde	5.67
2010 Large	26.05be	4.29
1930 Small	26.28bcde	6.86

7.3.2 NTR and FTSW

No significant difference ($p>0.05$) in the average transpiration rate between leaf size was found. There was a significant difference ($p<0.05$) in the average transpiration rate between the decade of cultivar release. Over the decades, the average transpiration rate increased from 1920 (705.89) to 1940 (1077.97) (Figure 7.2) before a decrease to 1960 (769.53). There was a slight increase in the 1970s (786.08) but a decrease to 1980 (690.32). The average transpiration rate increased in 1990 (767.55) followed by a slight decrease in 2000 (742.90) and an increase in the 2010s (843.67).

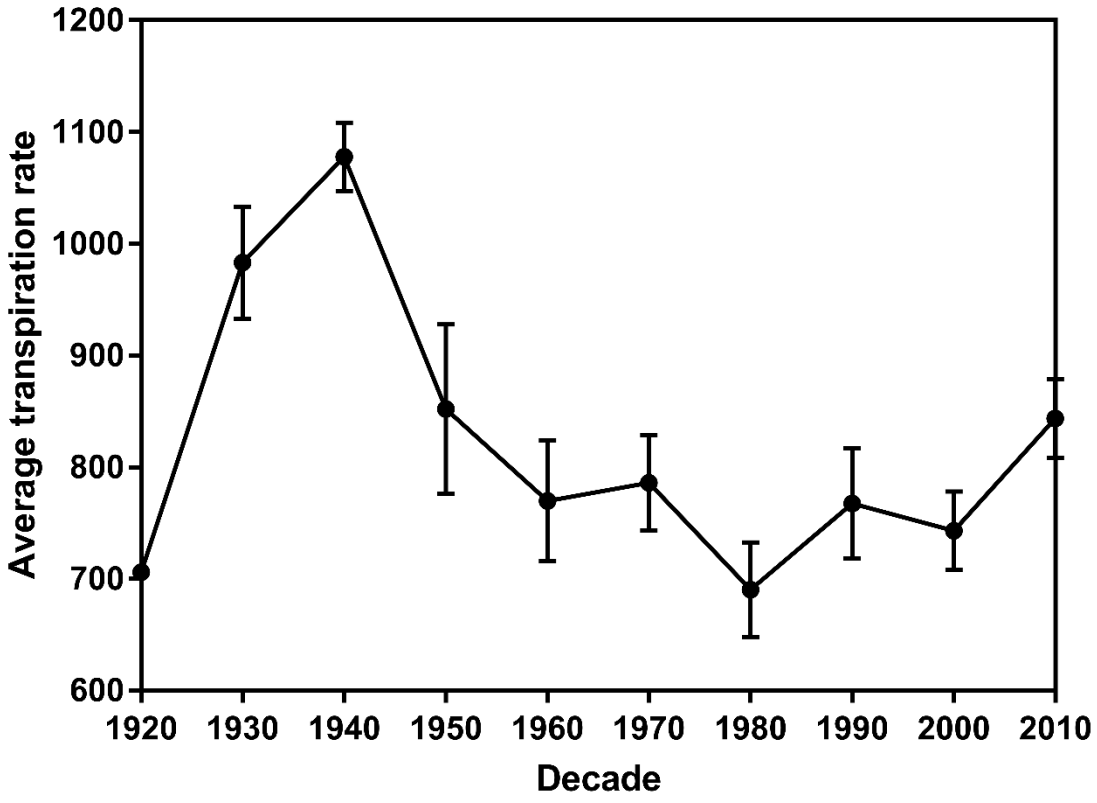

Figure 7.2 The average transpiration rates per decade of release for eighty white clover cultivars. The error bars are the standard error of the means.

A consistent relationship was found between NTR and FTSW values for each cultivar (Figure 7.3a and 7.3b). The pattern was similar to previous reports in other crops (Muchow, et al., 1991). The FTSW values ranged from 0.11 to 0.50, with an average of 0.29. On average, the NTR value was equivalent to well-watered plants until the FTSWc reached 0.29. The NTR value decreased linearly to 0 below an FTSW value of 0.29. Generally, there was no decrease in NTR until FTSW reached 0.7. The average number of days till the end of transpiration for the NTR was 9.95 and 14.48 for the FTSW.

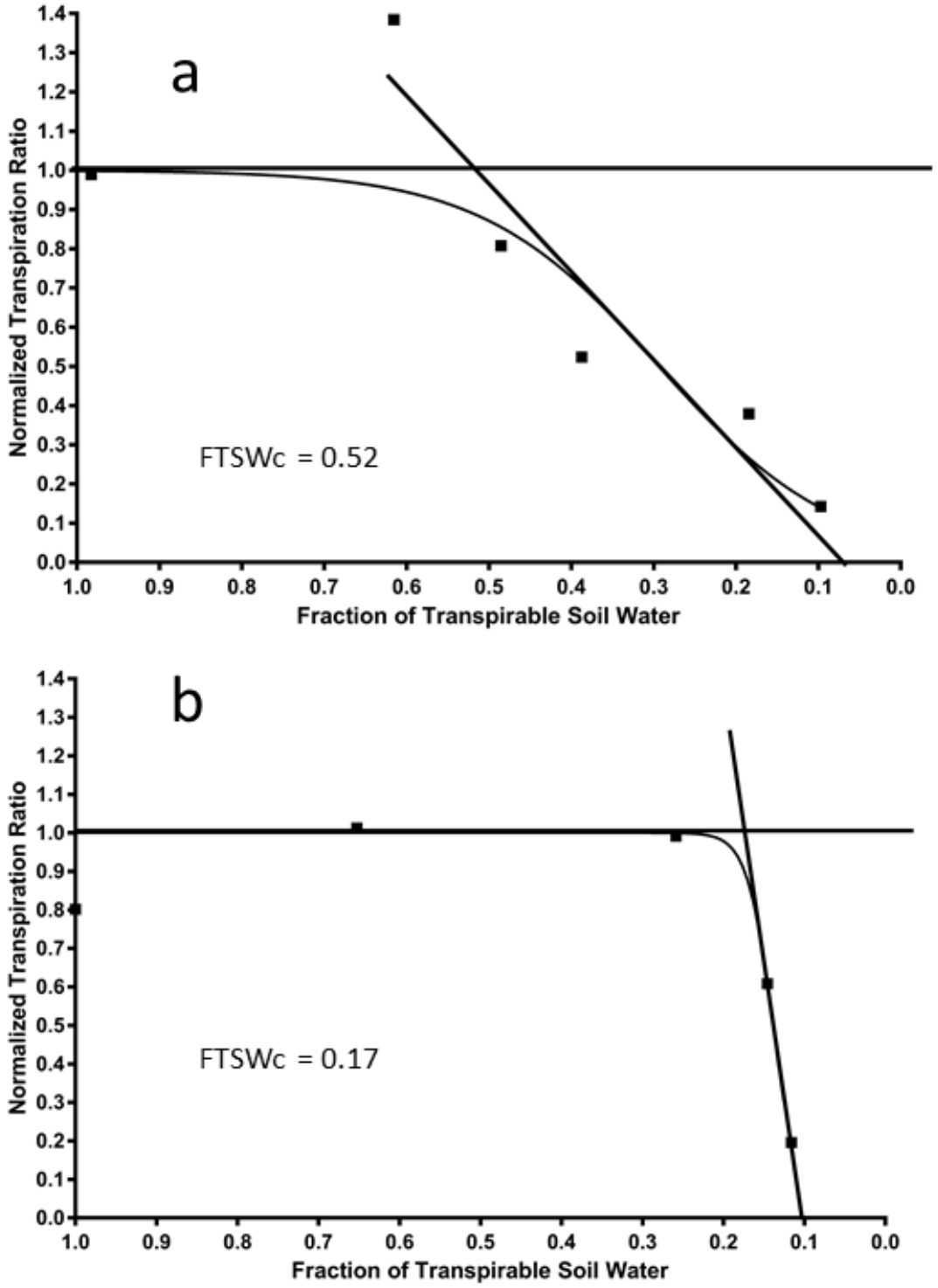

Figure 7.3 The average daily normalised transpiration ratio (NTR) response to the fraction of transpirable soil water (FTSW) for two white clover cultivars, (a) Tribute, released in 2000, and (b) Chieftain, released in 2000, that show contrasting FTSW thresholds. The FTSW threshold (FSTWc), where NTR begins to decrease, is shown.

7.3.3 Stomatal closure

Plateau regression was used to determine the FTSW value where the stomata permanently closed. The average FTSWc across the decades had an increasing trend from 1920 (0.23) to 1960 (0.38) (Figure 7.4). There was a slight decrease in the 1970s (0.34) before it increased to 0.38 in the 1980s. From the

1980s to 2010, there is a general decreasing trend until 2010 (0.23). The trend in the decade of release and FTSWc had an r^2 value of 0.81. The *P*-value of 0.016 associated with parameter *b* indicated that the FTSWc value significantly increased on average at a rate of 0.0329 (with a standard error of 0.013) per decade until 1960 (Table 7.6). Then, the *P*-value of 0.002 associated with parameter *c* indicated that the increasing trend changed significantly downwards after 1960. The *P*-value of 0.001 associated with the sum of parameters *b* + *c*, indicated that the FTSWc value significantly decreased on average at a rate of 0.0307 (with a standard error of 0.0089) per decade after 1960. There were statistical differences in FTSWc when the cultivars were grouped by decade and leaf size (Table 7.7).

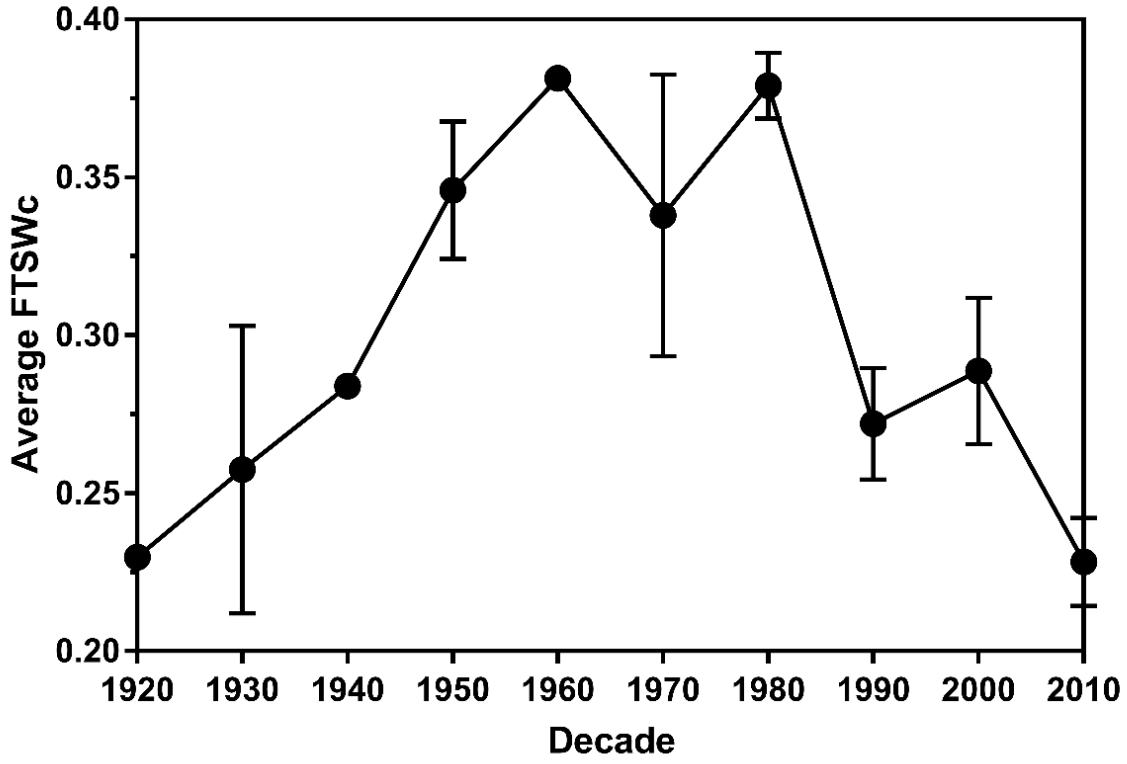

Figure 7.4 The average critical fraction of transpirable soil water threshold (FTSWc) per decade of release for eighty white clover cultivars. The error bars are the standard error of the means.

Table 7.6 The estimated linear spline trend: Inflection point = *a* + *b* * Time + *c* * Time_2, estimates and standard error (SE) of estimates and significance levels (*P*-values) of the estimates. *S* indicates a statistically significant estimate at 5% significance level.

Parameter	Estimate	SE	*P*-value	
a	0.2311	0.0464	< 0.001	***S***
b (= increasing rate per decade till 1960)	0.0329	0.0133	0.016	***S***
c	–0.0636	0.0199	0.002	***S***
b + *c* (absolute value = decreasing rate per decade after 1960)	–0.0307	0.0089	0.001	***S***

Table 7.7 The average FTSW critical threshold (FTSWc) for eighty white clover cultivars, clustered by leaf size and decade of release. Duncan's lettering compared the dry weights for leaf size and decade of release and the groups without a common small letter are significantly different at the 5% level of probability.

Decade of release and leaf size	FTSWc average	$LSD_{0.05}$
1930 Small	0.17a	0.23
2010 Medium	0.23ab	0.15
1920 Medium	0.23abc	0.22
2010 Large	0.24abc	0.16
1990 Medium	0.26abcd	0.13
2000 Large	0.26abcde	0.14
1940 Medium	0.28abcde	0.23
2000 Medium	0.30abcde	0.13
1960 Medium	0.30abcde	0.18
1930 Medium	0.30abcde	0.20
1990 Large	0.31abcde	0.15
1970 Large	0.31abcde	0.17
1950 Large	0.33abcde	0.23
1950 Medium	0.35abcde	0.17
1970 Medium	0.37acde	0.16
1980 Medium	0.38ce	0.14
1980 Large	0.40abcde	0.22

There was a significant ($p<0.001$) difference between the ten cultivars with the highest and lowest FTSWc (Figure 7.5). The five cultivars with the lowest FTSW were AberHerald (0.12), Aquiles (0.13), Kent White (0.17), Goliath (0.17) and Chieftain (0.17). The five cultivars with the highest FTSW were Kotare (0.44), Bounty (0.47), Pitau (0.48), Sacramento (0.49) and Tribute (0.52).

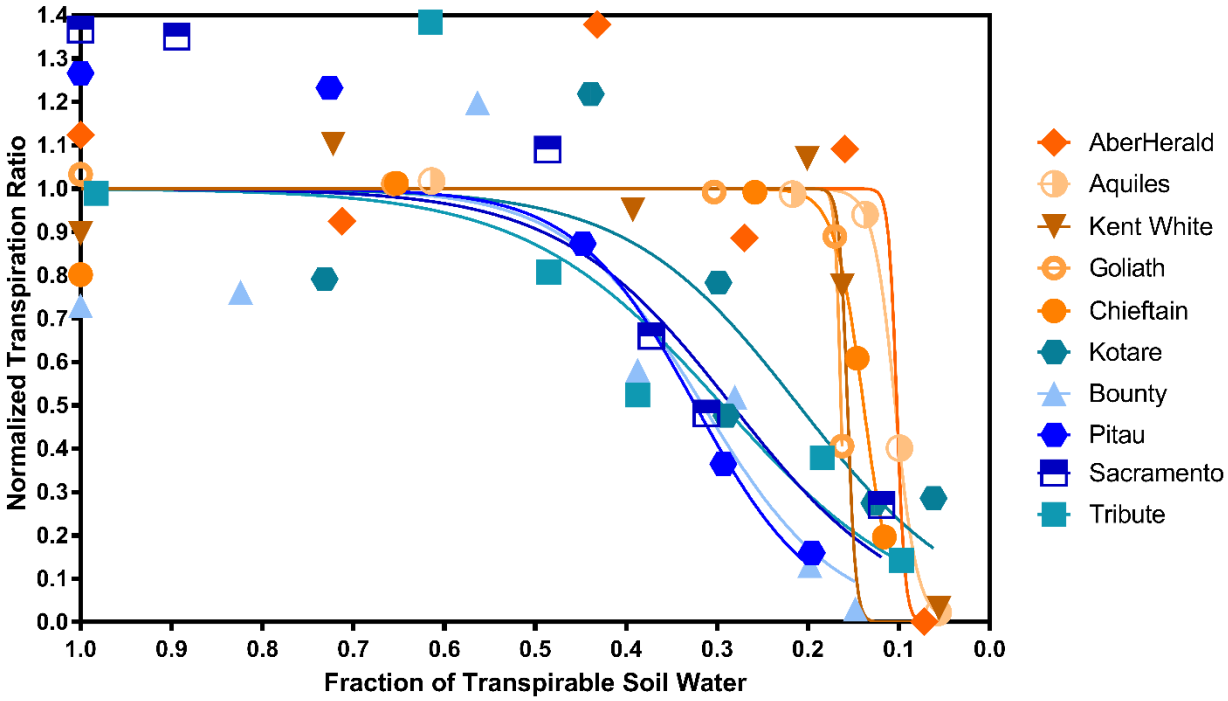

Figure 7.5 The five highest and five lowest critical fraction of transpirable soil water (FTSWc) thresholds for eighty white clover cultivars. The cultivars with the highest FTSW threshold (Kotare, Bounty, Pitau, Sacramento and Tribute) are denoted in blue, while the cultivars with the lowest FTSW threshold (AberHerald, Aquiles, Kent White, Goliath and Chieftain) are denoted in orange.

7.4 Discussion

7.4.1 FTSW threshold

The objective of this study was to evaluate the response of 80 white cultivars to drying soil. The results of this experiment are similar to previous reports. The transpiration rate in plant species has shown to be unaffected by drying soil until the FTSWc decreases to between approximately 0.25 to 0.35 (Gholipoor, et al., 2013, Lecoeur, et al., 1996, Meyer, et al., 1981, Miller, 2000, Ray, et al., 1997, Ray, et al., 1998, Ritchie, 1973, Rosenthal, et al., 1987, Sinclair, et al., 1986). The dependency of NTR and FTSW is shown through similar patterns of the NTR-FTSW relationship. The fact that the average FTSW in 1920 was similar to the average FTSW in 2010 suggests that breeding efforts have not implicated the relationship. Regardless of the decade of the release of the cultivar, the response to drought is similar.

White clover breeding in the 1900-1950s is documented most thoroughly in New Zealand compared to other countries such as Europe and the United States (Zeven, 1991). The large majority of the understanding of white clover genetics and diversity and the effect of selection techniques began in the mid-1960s and onwards (Williams, 1987, Williams, et al., 2007). Hoyos-Villegas, et al. (2019) showed that breeding progress of white clover cultivars could be divided into two eras; pre- and post-1965. There were significant increases in white clover sward content and dry matter yield after 1965,

but not pre-1965. The results in this study were consistent with the results found by Hoyos-Villegas, et al. (2019) as the significant increase and decrease of FTSWc could be divided into the same two eras; pre- and-post 1965.

Prior to 1965, the breeding decisions for white clover breeding programmes were based on increasing the performance of ecotypes and existing cultivars, and simple phenotypic selection. The breeding programmes and trials were performed across multiple regions and trial sites, and populations were selected for broad adaptation across a range of farming systems (Williams, et al., 2007). The breeding programmes relied on the variation that was present within countries and local environments, as there was little germplasm exchange between countries. Cultivars bred post-1965 utilized foreign germplasm and selection techniques such as recurrent phenotypic selection and wide hybridization (Ellis, et al., 1967, Williams, 2014). In the 1980s in New Zealand, a large and stable clover seed export market had been established. The cultivars that were exported were bred through a variety of techniques. Local ecotypes and local populations were utilized for adaptation to the target environment, and elite breeding populations were incorporated for a range of desirable traits. The populations were combined and evaluated through phenotypic selection methods (Caradus, et al., 1998a). The introduction of mixed swards to evaluate populations and grazing animals as a selection pressure occurred after 1965 (Woodfield, et al., 1994). It is possible that the breeding objectives of production-based traits took priority and tied with the intensification of agriculture, reduced the drought tolerance of germplasm.

Globally, germplasm exchange increased to widen the genetic base of populations in the later 1900s. Egan, et al. (2019a) and Egan, et al. (2019b) showed that in white and red clover, the introduction of foreign germplasm into the MFGC peaked in the 1970s and 1980s. By the 1980s, large-scale multi-country breeding programmes were established (Williams, et al., 2007). It is estimated that exotic germplasm used since the mid-1960s has contributed ~$1 billion annually to the pastoral agricultural exports (Lancashire, 2006). The characterization of the germplasm and new methods of utilizing the foreign germplasm could be expected to increase the performance of white clover germplasm (Egan, et al., 2019a, Williams, et al., 2007).

Although foreign germplasm has been utilized effectively worldwide (Rumball, et al., 1974), the world checklists of white clover cultivars show that the large majority of the cultivars with good tolerance to drought and heat were released before 1965. (Caradus, 1986, Caradus, et al., 1997b). The rankings could suggest that locally adapted germplasm has outperformed foreign germplasm for drought tolerance. However, germplasm collected from the Mediterranean has been used in breeding programmes to produce cultivars with increased winter-growth activity (Ayres, et al., 2007, Cooper, et al., 1997, Woodfield, et al., 2001b).

Throughout the decades of breeding, different breeding goals have been the focus of breeding programmes. Caradus, et al. (1989a) summarised the breeding goals for each decade of white clover

breeding in New Zealand. The early programmes focussed on advancing ecotypes and existing cultivars. In comparison, the later breeding programmes focussed on whole plant production and the integration of different farm and grazing management practices. The only decade to focus specifically on physiological and morphological responses to environmental changes was the 1950s, where it is likely that drought tolerance was integrated into cultivars released. A focal breeding target in the 1970s was on the production of white clover, and this benefited the increase in intensive farming in the 1980-1990s. The 1990s had many studies that focussed on the relationship between fertilizer application and grazing regimes and resulted in economic benefit (O'Connor et al. 1989; Roberts et al. 1992; Clark & Harris 1996). Selections for specific environments (Cooper, et al., 1993, Widdup, et al., 1989) and farming systems (Van den Bosch, et al., 1986, Williams, et al., 1979) have added value to New Zealand and the overseas market.

The lack of statistical significance difference between the average FTSW and the leaf size of cultivars suggests that all cultivars perform similarly under drought conditions. However, these results suggest that certain cultivars may be better utilized in certain environments and farming systems, i.e. short- and long-term droughts. Although the differences between all cultivars are not deemed statistically significant, the implications of a small difference in the NTR-FTSW relationship and FTSWc can be important knowledge for field conditions.

A low FTSWc suggests that the cultivar can sustain normal transpiration for a longer period in soils with less available water implying that it would perform advantageously under short-period drought conditions, compared to cultivars with a high FTSW (Ray, et al., 1997). AberHerald had the lowest FTSWc of the eighty cultivars. AberHerald is a medium leaved cultivar originating from Wales, United Kingdom. It performs well in cold environments, ensuring good stolon survival over winter. Helgadóttir, et al. (2001) showed that AberHerald showed morphological adaptation to more marginal climates. AberHerald has good stolon survival and rapid stolon recovery after grazing (Rhodes, et al., 1993). Other cultivars with a low FTSWc have high persistence under grazing (Charlton, et al., 1983, Widdup, et al., 2006).

A high FTSWc proposes that the cultivar can perform better in long-period drought conditions as they conserve water stores by closing their stomata early (Ray, et al., 1997). Tribute had the highest FTSWc of the eighty cultivars. Tribute is a medium-leaved cultivar bred from germplasm from New Zealand and Europe. Tribute was initially bred through a breeding programme for increased drought tolerance in Australia. Woodfield, et al. (2003) noted that Tribute had good drought performance in the third year of a grazing trial in Canterbury. Tribute produces a greater number of stolons and stolon density than other cultivars of medium leaf size (Widdup, et al., 2011).

7.4.2 Drought tolerance in white clover

This experiment gives an objective measurement of white clover cultivar transpiration rates and the critical FTSWc at which each cultivar begins to have permanent stomatal closure. This deepens the understanding of the expected response of cultivars to drought-like conditions.

The effect of drought conditions on plant growth and morphology is variable in white clover and is dependent on the length and intensity of the drought, and the stage of plant growth. Vegetative development is more sensitive than reproductive development (Belaygue, et al., 1996). While white clover can experience an increase in viable florets per inflorescence (Bissuel-Belaygue, et al., 2002a) and nectar amount and quality (Bissuel-Belaygue, et al., 2002b), leaf expansion and stolon branching characteristics are reduced during the development of each leaf in response to drought (Belaygue, et al., 1996). Belaygue, et al. (1996) suggests that branching and leaf appearance on the main stem are, respectively, the most and least sensitive processes to drought conditions. However, transpiration is more sensitive than photosynthesis and reproductive development. Transpiration is reduced in relation to stomatal closure (Lacape, et al., 1998, Lecoeur, et al., 1996). Nonetheless, while a range of drought conditions can close stomata and reduce transpiration, net photosynthesis is kept at a maximum level (Wery, 2005).

The breeding of the *Trifolium* ISH programme has created hybrids that have increased drought tolerance compared to white clover (Marshall, et al., 2001). *T. ambiguum*, *T. uniflorum* and *T. occidentale* are the related species that have been hybridized with white clover. The programme aims to integrate the drought-tolerant traits from the related species into white clover. Nichols, et al. (2014c) demonstrated that the *T. repens* x *T. uniflorum* ISH were more drought-resistant than white clover. The first generation BC_1 had less of a decrease in stolon parameters and senescence than in BC_2 hybrids and white clover. It was hypothesized that the compact growth form of the BC_1 hybrids allowed increased allocation of dry matter to the roots under drought. In a similar study, Nichols, et al. (2013c) found that the net photosynthetic rate was decreased by up to 48% in BC_2 hybrids and white clover, but there was no significant decrease for the BC_1 hybrids. They concluded that due to the effects on stolon morphology, the BC_1 hybrids might be able to maintain a higher water uptake through increased allocation to root biomass than white clover during drought conditions.

7.4.3 Canopy wilting

The underlying mechanisms of drought-tolerant phenotypes are many. Canopy wilting is one of the first signs of drought stress caused by soil water deficits (Kunert, et al., 2020). In soybean, the slow-wilting phenotype was first reported in a Japanese landrace (Sloane, et al., 1990) and the development of slow-wilting genotypes has enabled selection for breeding (Sinclair, et al., 2010). The development of genotypes with delayed canopy wilting phenotypes have been studied thoroughly in soybean and in several plant introductions, and could lead to increased yield stability in drought conditions (Steketee,

et al., 2020). Simulations have suggested if the phenotype was bred into populations, yield in drought conditions could improve by >80% (Sinclair, et al., 2010). However, the physiological mechanisms controlling the slow-wilting phenotype remain uncertain. Failure to understand the mechanisms will constrain breeding efforts.

Several traits appear to be involved in a drought-tolerant phenotype, including a large root system (Pantalone, et al., 1996) and constant transpiration rates under water vapour deficit (Fletcher, et al., 2007). Hofmann, et al. (2000) suggested that white clover populations that have adapted to marginal climates have an increased ability to accumulate ortho-dihydroxylated flavonoids. The populations showed increased tolerance to UV-B radiation. Mechanisms involved in the slow-wilting phenotype may include the accumulation of minerals (Bellaloui, et al., 2013) and factors that affect transpiration, such as the distribution and expression of aquaporins (Fletcher, et al., 2007, Sinclair, et al., 2008). Furthermore, Ye, et al. (2020) compared two slow-wilting soybean accessions. Although both had limited-maximum transpiration rates, transpiration was sensitive to an aquaporin inhibitor, supporting the hypothesis that multiple water-conservation mechanisms are involved in the slow-wilting phenotype. They suggested that soybean genotypes with the slow-wilting phenotype restrict transpiration through decreased radiation use efficiency and improved water use efficiency. Sadok, et al. (2010) demonstrated that the slow-wilting phenotype exhibited a limited leaf hydraulic conductance for transpiration rate under high VPD, indicating that the symplastic pathway terminating in the guard cells of the leaves was involved in delaying wilting. In addition, Bagherzadi, et al. (2017) showed that the common mechanisms controlling plant water transport were not involved in the slow-wilting trait and suggested that anatomical features associated with precise water transport characteristics could be involved.

Recent studies have increased our understanding of the underlying genetic architecture of the slow-wilting phenotype. QTL have been identified for canopy wilting (Abdel-Haleem, et al., 2012, Ye, et al., 2020) and it has been concluded that it is a polygenic trait (Charlson, et al., 2009). A recent study by Kaler, et al. (2017) has identified SNPs associated with canopy-wilting that are located within or close to genes with connections to transpiration or water transport. Devi, et al. (2015) analysed gene expression in the leaves of two slow-wilting accessions and showed that 944 genes were differentially expressed in one accession compared to the other. While more recently, Steketee, et al. (2020) used a GWAS to identify 45 marker-trait associations with canopy wilting and as a result, several new accessions were identified with the slow-wilting phenotype. Although the research into the slow-wilting phenotype has primarily been performed in soybean, other crops, such as cowpea (*Vigna unguiculata* (L.) Walp) have utilized the phenotype to identify accessions with increased drought tolerance (Pungulani, 2014).

The utilization of a slow-wilting phenotype in white clover could increase the performance under drought conditions. Although the cultivars in this study have been characterized by FTSWc under

drought conditions, germplasm exploration to identify accessions with a slow-wilting phenotype is needed to accelerate breeding efforts (Barbour, et al., 1996).

7.4.4 Conclusions

The results from this study highlight the variable rates of stomatal closure for eighty white clover cultivars. The relationship between NTR and FTSW is consistent for all cultivars, regardless of the decade of release. Cultivars that have a significantly higher or lower FTSWc have been identified and have deepened the knowledge of the cultivar response to drought conditions. The white clover ISH programme shows promise for increasing drought tolerance but further replicated trials are needed to assess performance. The increasing demand for cultivars to perform under extreme conditions in response to climate change requires more research into the genetic and phenotypic basis of drought traits and how these can be incorporated into breeding programmes.

Chapter 8

Genome-wide association study of agronomic traits in white clover

8.1 Introduction

A white clover cultivar must be high performing in several agronomic traits to be successful. As mentioned in chapter 6, herbage yield in a sward is a critical breeding target. The amount of clover present in a sward, termed clover content, is vital as white clover grows in companionship with grass and must compete but remain balanced. Clover content affects the production of the grazing livestock, e.g. milk yield for dairy cows, and therefore affects farm profit (Harris, et al., 1997, Orr, et al., 1990). The improvement of seed traits such as seed yield and flowering characteristics are also important breeding objectives. There are differences between cultivars for seed traits, implying that there is genetic variability (Connolly, 1990, Williams, et al., 1998).

Drought stress is one of the most limiting factors in white clover production. Breeding for increased drought tolerance is a highly desirable characteristic (Knowles, et al., 2003). However, drought tolerance is an extremely complex trait that takes many breeding cycles to develop an elite performing population. As discussed in chapter 7, measurements of NTR and FTSW can infer drought tolerance.

Unlike drought tolerance, cyanogenesis, the production of hydrogen cyanide (HCN), is a simple trait. White clover is naturally polymorphic for cyanogenesis. The production of HCN acts as a plant defence response to herbivores (Dirzo, et al., 1982, Pederson, et al., 1998). However, this is often bred to null or low levels early-on in breeding programmes as high levels of HCN can be detrimental to animal health (Bishop, et al., 1969). The hydrolysis of cyanogenic glucosides generates HCN. Two genes control cyanogenesis; the *Ac/ac* allele controls the presence or absence of cyanogenic glucosides in the vacuoles of the leaf and stem tissue. The *Li/li* allele controls the presence or absence of linamarase in the cell wall. Linamarase influences the rate of enzymatic breakdown of the glucosides. The *Li* allele is located on a different chromosome to the glucoside locus. Both *Ac* and *Li* show incomplete dominance. Two non-functional alleles at either gene produce an acyanogenic phenotype (Hughes, 1991).

Leaf size is a simple trait in white clover and is used to classify white clover cultivars; large, medium or small. The characteristics and production systems that the different leaf sized cultivars are used in are explained in chapter 2.2.4.

Utilising molecular tools could increase the rate of genetic gain in white clover. However, the highly outcrossing and heterozygous nature of white clover has challenged the development of molecular tools. Early work in the advancement of white clover genetics began with the construction of linkage maps (Barrett, et al., 2004, Griffiths, et al., 2013, Jones, et al., 2003, Zhang, et al., 2007) and the

identification of QTL for seed yield traits (Barrett, et al., 2005, Barrett, et al., 2009), morphological traits such as petiole length and plant height (Cogan, et al., 2006), and forage yield (Jahufer, et al., 2013b). Studies also focussed on the identifying genetic variation in white clover that could be utilised in breeding programmes (Gustine, et al., 2002, Jahufer, et al., 2008). Recent advances in genomics have accelerated breeding in forages (Bouton, 2010, Brummer, 2013). The identification of markers that are tightly linked to regions of the genome that control important agronomic traits could increase the genetic gain and speed of selection cycles.

GWASs are a powerful method to detect genomic associations with traits of interest in unstructured populations (Sebastiani, et al., 2009, Song, et al., 2015). There is currently only one GWAS published in white clover (Inostroza, et al., 2018). GBS is a cost-effective, high-throughput, and increasingly popular method to discover SNPs (D'Agostino, et al., 2018, Huang, et al., 2014). GBS is suitable for species with large, diverse genomes and is useful for population studies, germplasm characterisation, breeding and mapping traits (Elshire, et al., 2011). Studies have used GBS to generate large numbers of SNPs to investigate species diversity, genomic prediction and genome-wide association studies (GWAS) (Biazzi, et al., 2017, Han, et al., 2018, Li, et al., 2015, Sakiroglu, et al., 2017, Sonah, et al., 2015). GBS produces adequate genome-wide markers to support GWAS. SNPs are a popular choice of marker to integrate into breeding programmes due to their low cost, and high abundance and high polymorphism rate throughout the genome (Acquaah, 2012). Recent developments in SNP markers in alfalfa (*Medicago sativa*) have been successful (Biazzi, et al., 2017, Sakiroglu, et al., 2017).

The objectives of this chapter were to use GBS SNP markers to (i) determine the population structure of the white clover Genetic Gain panel, and (ii) identify genomic regions controlling the traits of yield, clover content, seed yield, flowering duration, the peak number of flowers, leaf size, seed per head, leaf marking, cyanogenesis, normalised transpiration rate and the fraction of transpirable soil water.

8.2 Material and Methods

8.2.1 Plant material and phenotyping

The plant material was the same 80 cultivars as used in chapter 7. The same trial was used to collect the phenotypic data of dry matter yield, clover content, seed yield, flowering duration, the peak number of flowers, leaf size and seed per head. Hoyos-Villegas, et al. (2019) published the dry matter yield and clover content phenotypic data used in this GWAS. Briefly, the trials were planted in three locations (Canterbury, Manawatu and Waikato) in September 2014 and evaluations were performed until September 2017. One site (Manawatu) had to be replanted in autumn 2015 but remained grown for 3 years until September 2018. The trials were grazed under different grazing managements by sheep, dairy cattle and beef cattle as characterised by the region. Every season, measurements were taken every 4 to 6 weeks before grazing and biomass harvests were collected once, apart from spring

where two collections were collected. The R package *lme4* generated the BLUPs, using a mixed model to account for random effects.

All 80 white clover cultivars were tested for a cyanogenic phenotype. The picric acid test determined the presence of hydrogen cyanide (Nowosad, et al., 1940). One leaflet from each cultivar was placed in an Eppendorf with filter paper in the lid. 10μL of picric acid was added to the filter paper, and 10μL of toluene was added to each tube. The tubes were incubated at 37˚C for two hours. For each cyanogenic plant, a scale of intensity (1-4) was used; 1 as the least intense and 4 as the most intense.

Leaf marking characteristics were described as in (Tashiro, et al., 2010). The NTR and FTSW values calculated in chapter 7 were used to test for associations.

8.2.2 Genetic markers

DNA preparation and genotyping

The DNA extraction followed the protocol as described in Anderson, et al. (2018). GBS was performed as described by Elshire, et al. (2011) using the enzyme ApeK1. The initial sequencing effort of 80 samples with 5 replicates were sequenced on the Illumina Hiseq instrument compiled in 6 lanes. 1.2 billion reads were produced of 100bp; approximately 200 million reads per lane.

SNP calling and filtering

Raw reads were first demultiplexed using the GBSX program (Herten, et al., 2015). The Burrows-Wheel Aligner (Li, et al., 2009) was used to map the demultiplexed reads against the white clover reference genome. 'ref_map.pl' in the stacks pipeline version v 1.47 (Catchen, et al., 2013) was used with the parameters; minimum number of raw reads required to form a stack/putative allele (*m*) = 3, number of mismatches allowed between stacks/putative alleles to merge them into a putative locus (*M*) = 3, and number of mismatches allowed between stacks/putative loci during the construction of the catalogue (*n*) = 3. The stacks pipeline initially generated 234,367 SNPs. vcftools was used to filter the SNPs (Danecek, et al., 2011). The SNPs were filtered for a MAF <5%, markers with >80% missing data and a depth of <5 and >300. SNPs were not filtered for HWE as the panel was not a single population. The final set of SNPs contained a total of 69,202 SNPs.

8.2.3 Statistical analyses

Linear mixed models

The BLUPs for seed yield, flowering duration, the peak number of flowers, leaf size and seed per head were generated using the same methodology as for the yield and content BLUPs in Hoyos-Villegas, et al. (2019). The BLUPs were estimated by using the R package *lme4*, using a mixed model to account for random effects.

Analysis of variance

The cyanogenesis data were analysed using ANOVA in GenStat 18th edition (VSN-International, 2019). The mean, least significant difference (LSD) and F pr value were used as a test of significance. The replicate was used as the block and the plant cultivar used as the treatment. Data were normalised by log transformation in Minitab (Minitab, 2006).

Principal component analysis

The same methodology was used as in chapter 6.

8.2.4 Population structure and linkage disequilibrium

Linkage disequilibrium is defined as the non-random association of alleles at two or more loci (Slatkin, 2008). The estimation of LD in biallelic markers can be through the coefficient of correlation (r^2) (Kimura, et al., 1971). LD decay occurs because the LD between the loci and gene decays from recombination. The same methodology was used in chapter 6.

8.2.5 Genome-wide association analysis

The same methodology was used in this GWAS, as was used in chapter 6. However, the false discovery rate (FDR) was also used for multiple testing (Benjamini, et al., 1995). The R package, *qvalue*, was used to account for false positives by using FDR statistics (Dabney, 2010). The FDR was calculated by the formula (Benjamini, et al., 1995):

$$FDR = \mathbf{E}\left[\frac{V}{R \vee 1}\right] = \mathbf{E}\left[\frac{V}{R}|R > 0\right] \mathbf{P}(R > 0)$$

where, V is the number of type 1 errors, and R is the number of rejected hypotheses.

8.3 Results

8.3.1 Genetic markers

A total of 69,202 SNPs were present after filtering. The mean and minimum co-call rate for sample pairs was 0.70 and 0.45, respectively. The mean sample depth was 40.05, and the mean call rate was 0.82. The MAF values ranged from 0.05-0.5 (Figure 8.1). The SNP depth is visualised in Figure 8.2.

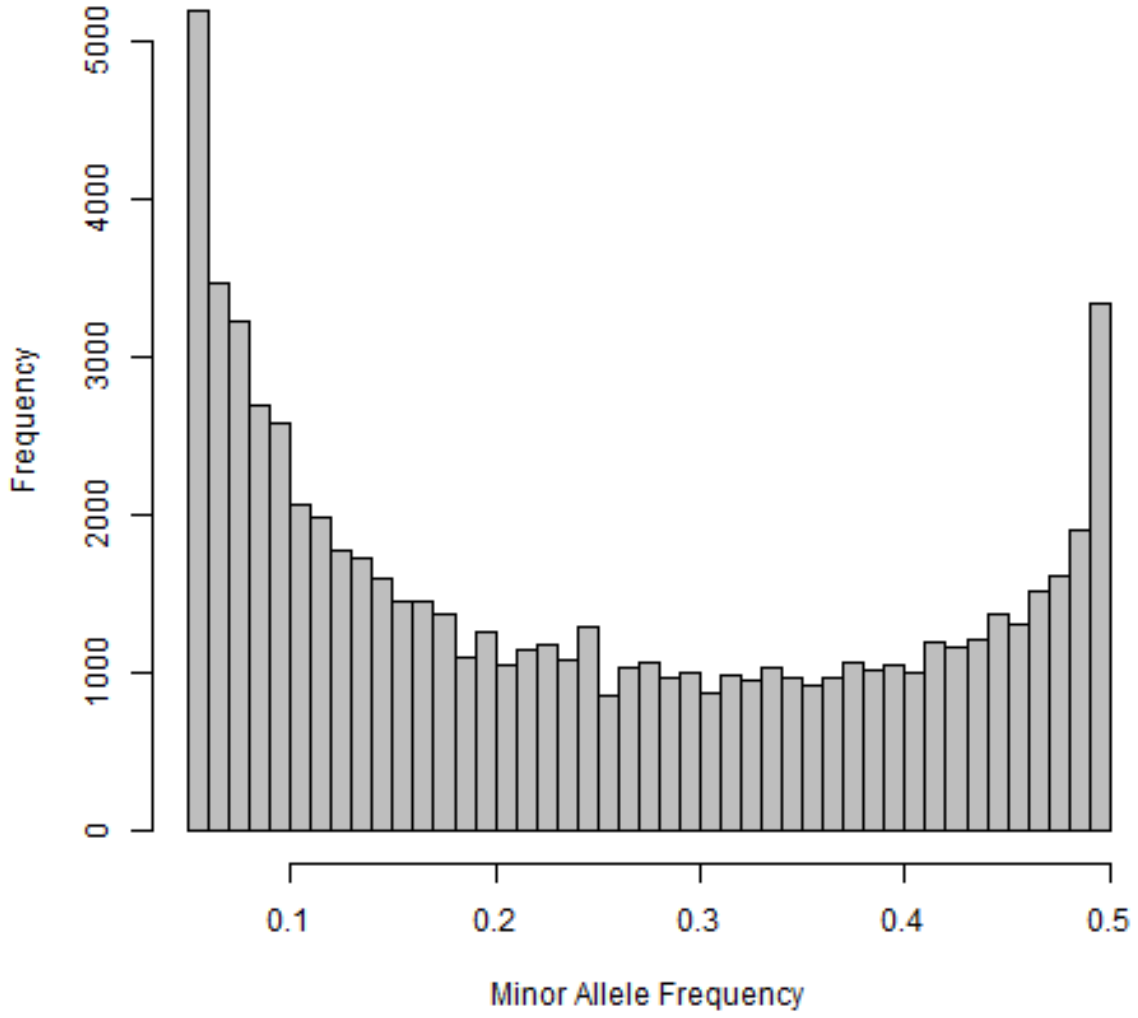

Figure 8.1 Minor allele frequency (MAF) for the 69,202 SNPs in the 80 white clover cultivars.

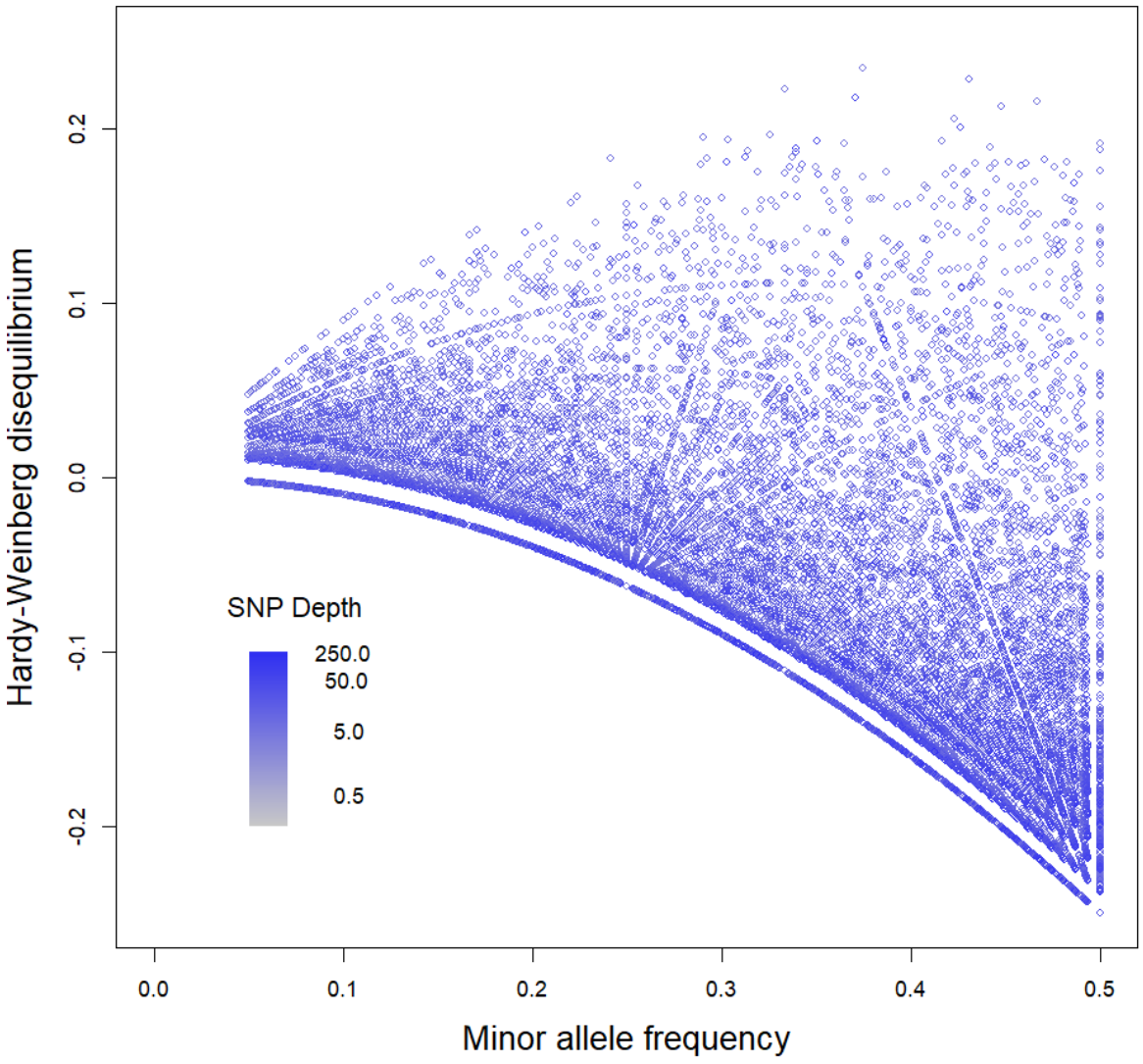

Figure 8.2 Finplot of the 69,202 filtered SNPs in the 80 white clover cultivars. The Hardy-Weinberg disequilibrium is plotted against the minor allele frequency (MAF) and shaded by the SNP depth.

8.3.2 Phenotypic data

ANOVA of the HCN scores showed a significant difference ($p<0.01$) within the panel (Figure 8.3). The HCN scores of the cultivars ranged from 0-4. The mean of the panel was 1.39 and had an $LSD_{0.05}$ of 0.79. When tested against the decade of release, the differences were significant ($p<0.05$). However, there were no significant differences ($p>0.05$) for HCN scores against the country of origin.

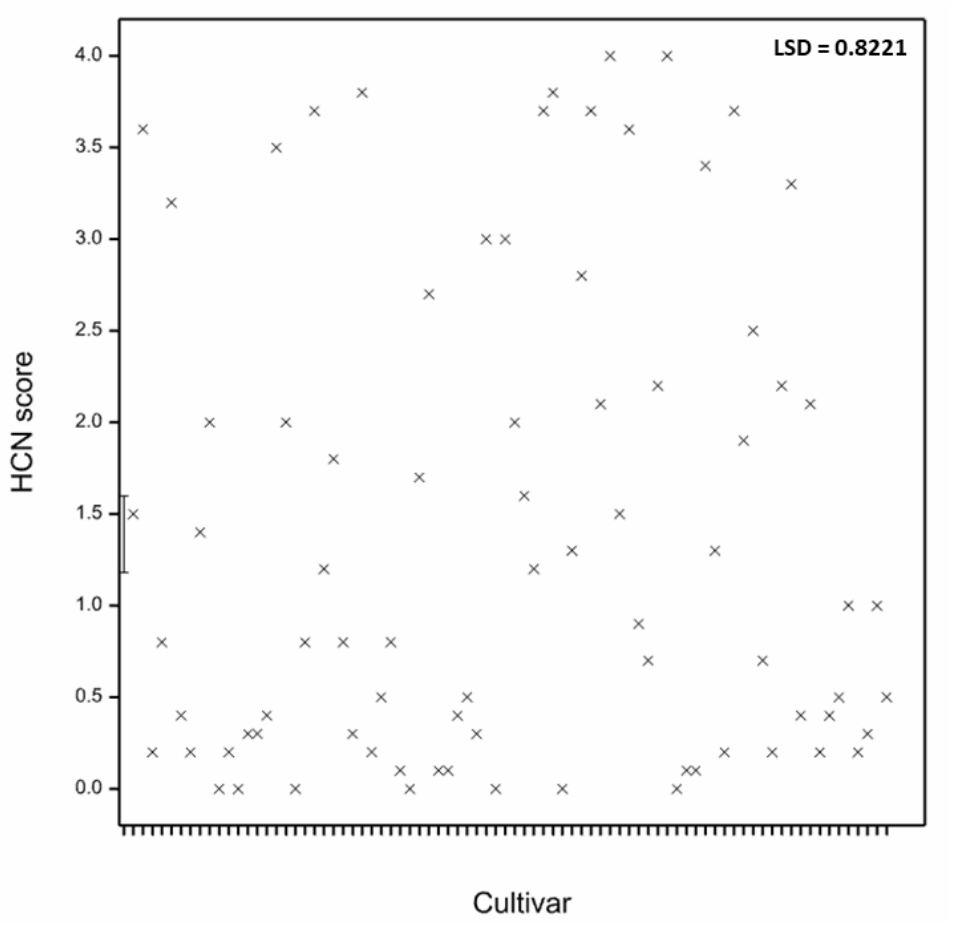

Figure 8.3 The mean HCN scores for 80 white clover cultivars. The LSD value was 0.8221. The error bar (0.4188) is the standard error of the mean.

The PC analysis of the 21 traits indicated that PC1, PC2 and PC3 explain 32.08%, 25.75% and 12.80% of the total variance, respectively. The PC analysis of dry matter yield, content, seed yield, flowering duration, the peak number of flowers, leaf size, seed per head, inflection point, HCN and leaf marking BLUPs generated a biplot (Figure 8.4) and a scree plot (Figure 8.5). Directional vectors indicate the correlation among the traits in the biplot. Symbols indicate the genotypes. PC1, PC2 and PC3 explained 35.27%, 28.70% and 14.20%, respectively, of the total variation. The biplot showed a positive association (angles between the directional vectors are less than 90 degrees) content, dry matter yield, seed per head, leaf size, HCN, inflection point, leaf marking and flowering duration. Positive associations were also found between seed yield and peak number of flowers. Negative associations were found between peak number of flowers and content, dry matter yield, seed per head, HCN, inflection point and leaf size. When clustered by country of origin and decade of release, there are no strong clustering patterns. The cultivars released in the 2000 and 2010 decades are clustering have a weak cluster (Figure 8.6).

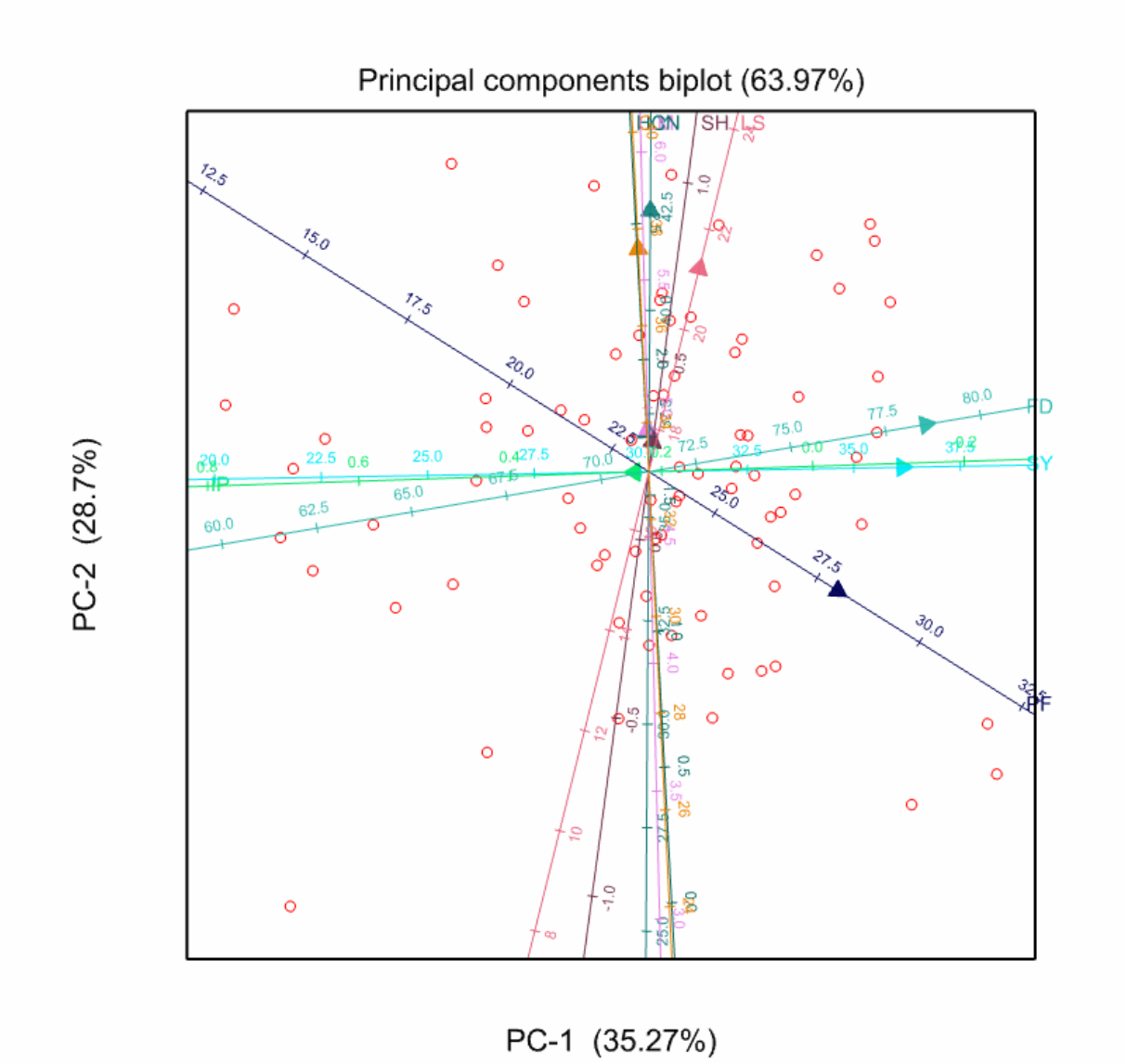

Figure 8.4 Biplot generated using standardised best linear unbiased predictor (BLUP)-adjusted means for seven traits measured from 80 white clover cultivars. PC1 accounted for 35.27% and PC2 28.70% of the total variation present. The traits are indicated by the directional vectors: dry matter yield (Y), content (C), seed yield (SY), flowering duration (FD), the peak number of flowers (PF), leaf size (LS), cyanogenesis (HCN) and seed per head (SH). Red circles indicate the cultivars.

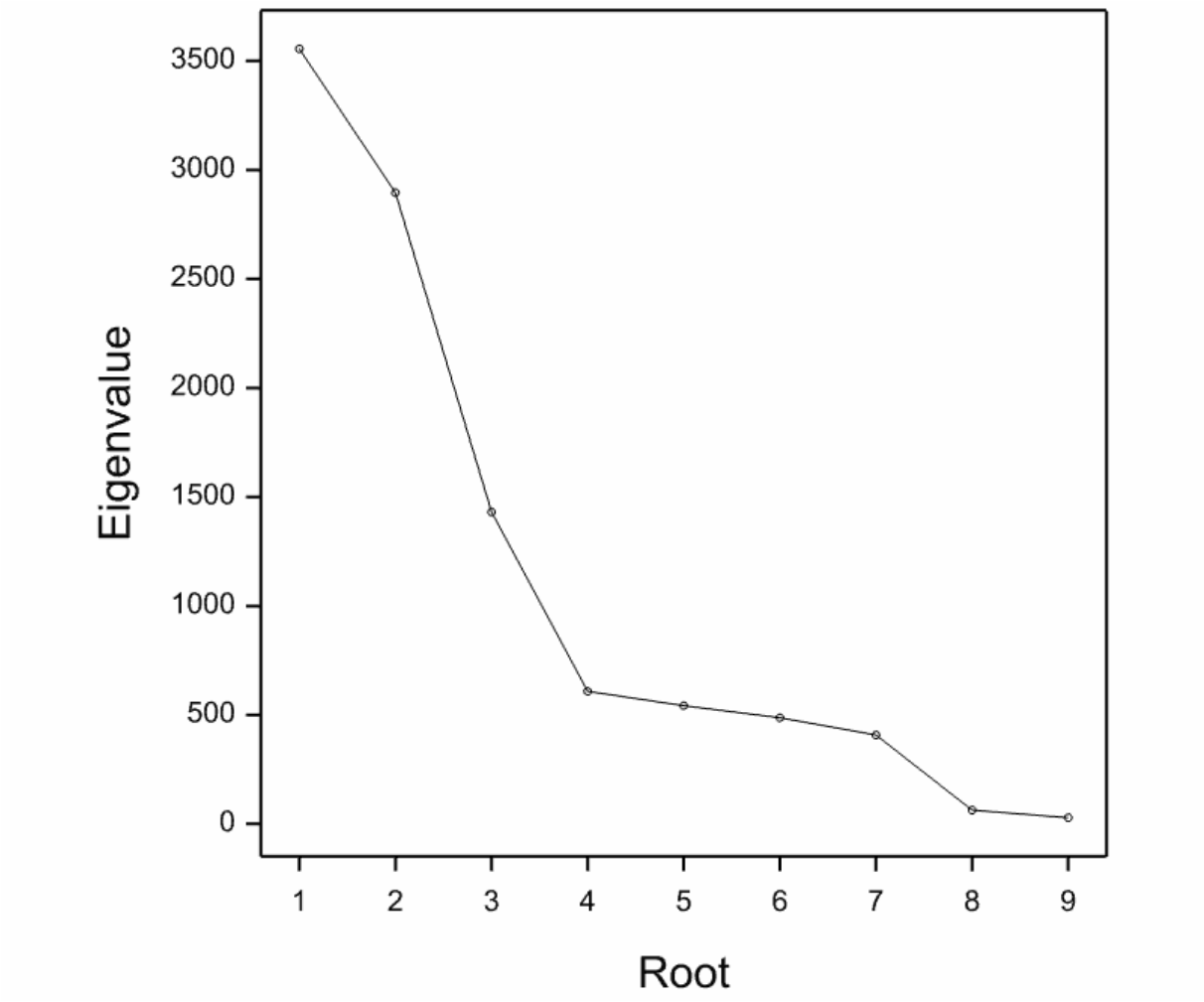

Figure 8.5 Scree plot generated using standardised best linear unbiased predictor (BLUP)-adjusted means for seven traits, dry matter yield, content, seed yield, flowering duration, the peak number of flowers, leaf size, cyanogenesis and seed per head measured from 80 white clover cultivars.

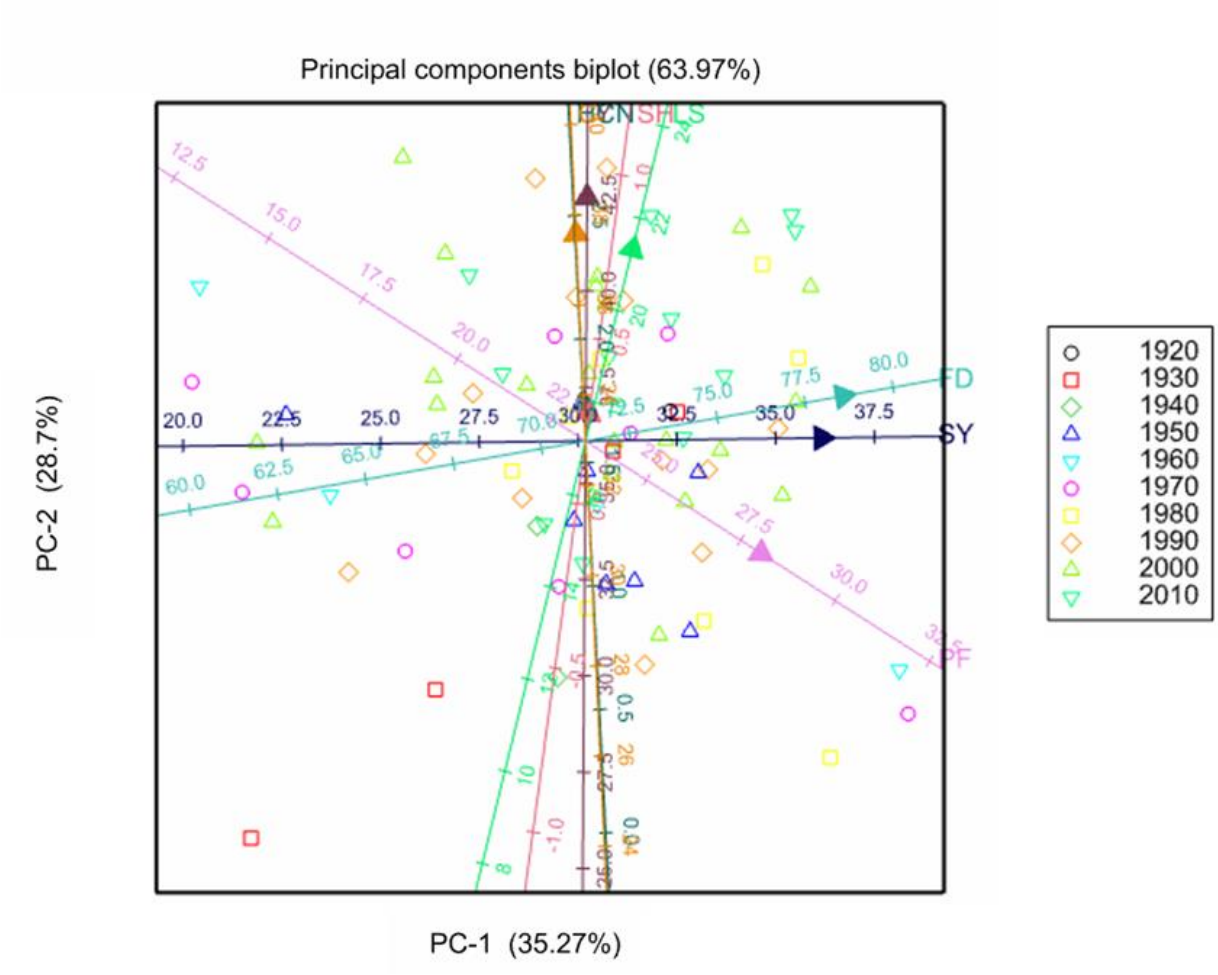

Figure 8.6 Biplot generated using standardised best linear unbiased predictor (BLUP)-adjusted means for seven traits measured from 80 white clover cultivars clustered by the decade of release. PC1 accounted for 35.27% and PC2 28.70% of the total variation present. The traits are indicated by the directional vectors: dry matter yield (Y), content (C), seed yield (SY), flowering duration (FD), the peak number of flowers (PF), leaf size (LS), cyanogenesis (HCN) and seed per head (SH).

Principal component analysis of the NTR and FTSW values showed that PC1, PC2 and PC3 explained 49.55%, 15.68% and 14.19% of the total variation, respectively (Figure 8.7 and Figure 8.8). Positive associations were found between all the FTSW days and NTR day 9. Positive associations were also found between FTSW days 3 and 5 and NTR days 5, 7 and 9. Seed yield, flowering duration, peak number of flowers, FTSW3, FTSW5, FTSW9, FTSW11, FTSW13, NTR3, NTR5, NTR7, NTR9 and FTSWc were significantly (p<0.05) associated with PC1 (Table 8.1). Dry matter yield, content, peak number of flowers, leaf size, seed per head, HCN, leaf marking, FTSW5, FTSW7, FTSW9, FTSW11, FTSW13, NTR3, NTR5, NTR7, NTR9 and FTSWc were significantly associated with PC2. There were no clear clustering patterns when grouped by the decade of release and country of origin.

Table 8.1 Correlation coefficients (*P*) and probabilities (*r*) for 21 traits for 80 white clover cultivars linked to the first two principal components (PCs). The means of the traits were regressed against the scores of PC1 and PC2.

	PC1		PC2	
Trait description	*P*	*r*	*P*	*r*
Dry matter yield	0.958	0.006	0.001	0.9094
Content	0.678	-0.0472	0.001	0.8974
Seed yield	0.001	0.8404	0.908	0.0132
Flowering duration	0.001	0.8719	0.236	0.1341
Peak number of flowers	0.001	0.7255	0.001	-0.4126
Leaf size (mm)	0.096	0.1875	0.001	0.6835
Seed per head (g)	0.581	0.0626	0.001	0.4165
HCN	0.902	-0.014	0.034	0.2367
Leaf marking	0.965	-0.0049	0.049	0.2204
FTSW day 3	0.001	0.8861	0.092	0.1894
FTSW day 5	0.001	0.9093	0.022	0.2567
FTSW day 7	0.001	0.9134	0.005	0.3143
FTSW day 9	0.001	0.8896	0.007	0.2985
FTSW day 11	0.001	0.8872	0.003	0.3267
FTSW day 13	0.001	0.7072	0.001	0.4484
NTR day 1	0.454	0.0849	0.07	-0.2038
NTR day 3	0.001	-0.4481	0.001	0.4605
NTR day 5	0.001	0.5738	0.001	-0.5224
NTR day 7	0.001	0.5439	0.001	-0.6521
NTR day 9	0.001	0.7522	0.001	-0.3705
FTSWc	0.001	0.4764	0.001	0.528

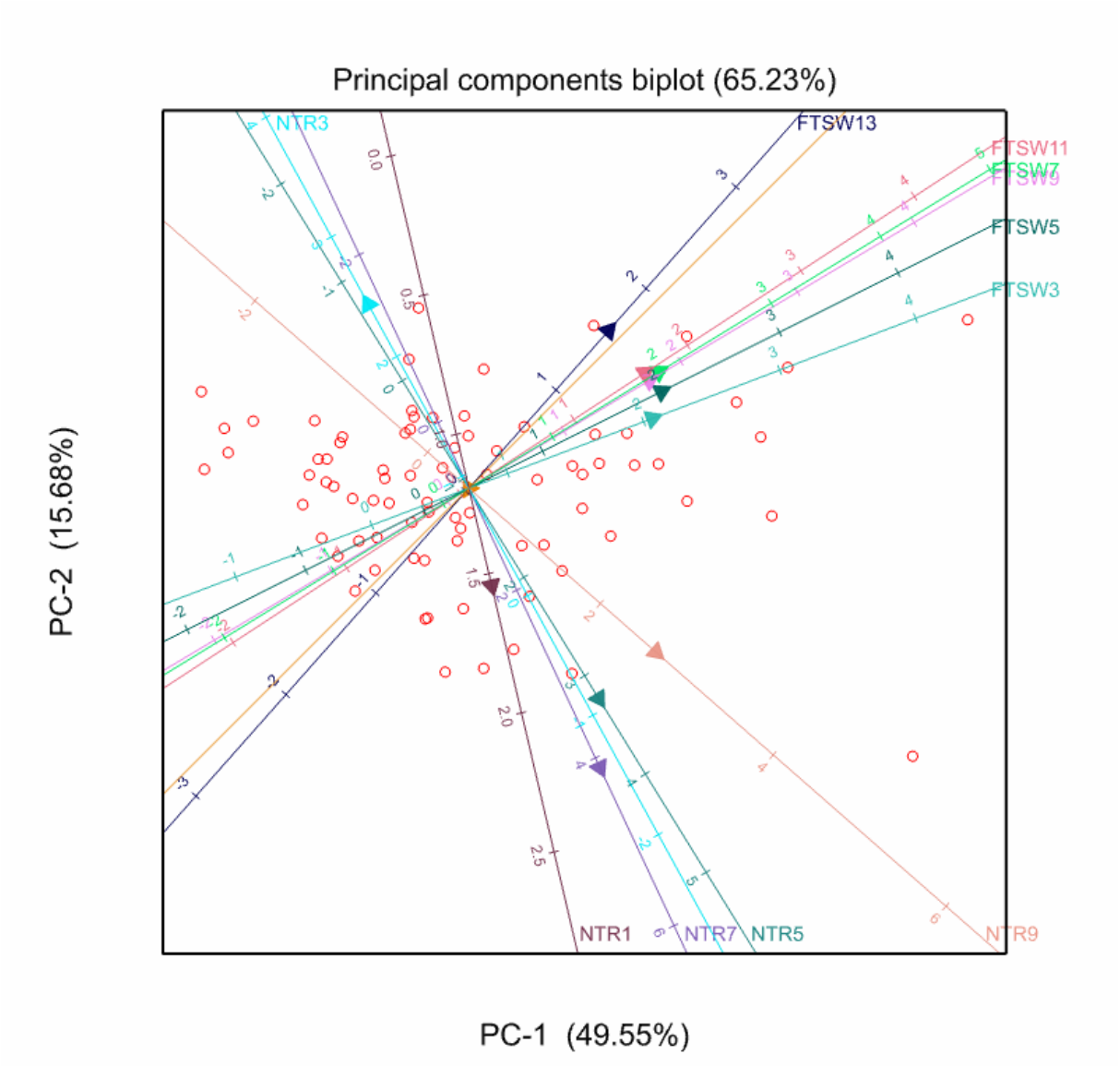

Figure 8.7 Biplot generated using standardised, normalised transpiration rate (NTR) and fraction of transpirable soil water (FTSW) values measured from 80 white clover cultivars. PC1 accounted for 49.55% and PC2 15.68% of the total variation present. The traits are indicated by the directional vectors: NTR day 1 (NTR1), NTR day 3 (NTR3), NTR day 5 (NTR5), NTR day 7 (NTR7), NTR day 9 (NTR9), FTSW day 3 (FTSW3), FTSW day 5 (FTSW5), FTSW day 7 (FTSW7), FTSW day 9 (FTSW9), FTSW day 11 (FTSW11) and FTSW day 13 (FTSW13). Red circles indicate the cultivars.

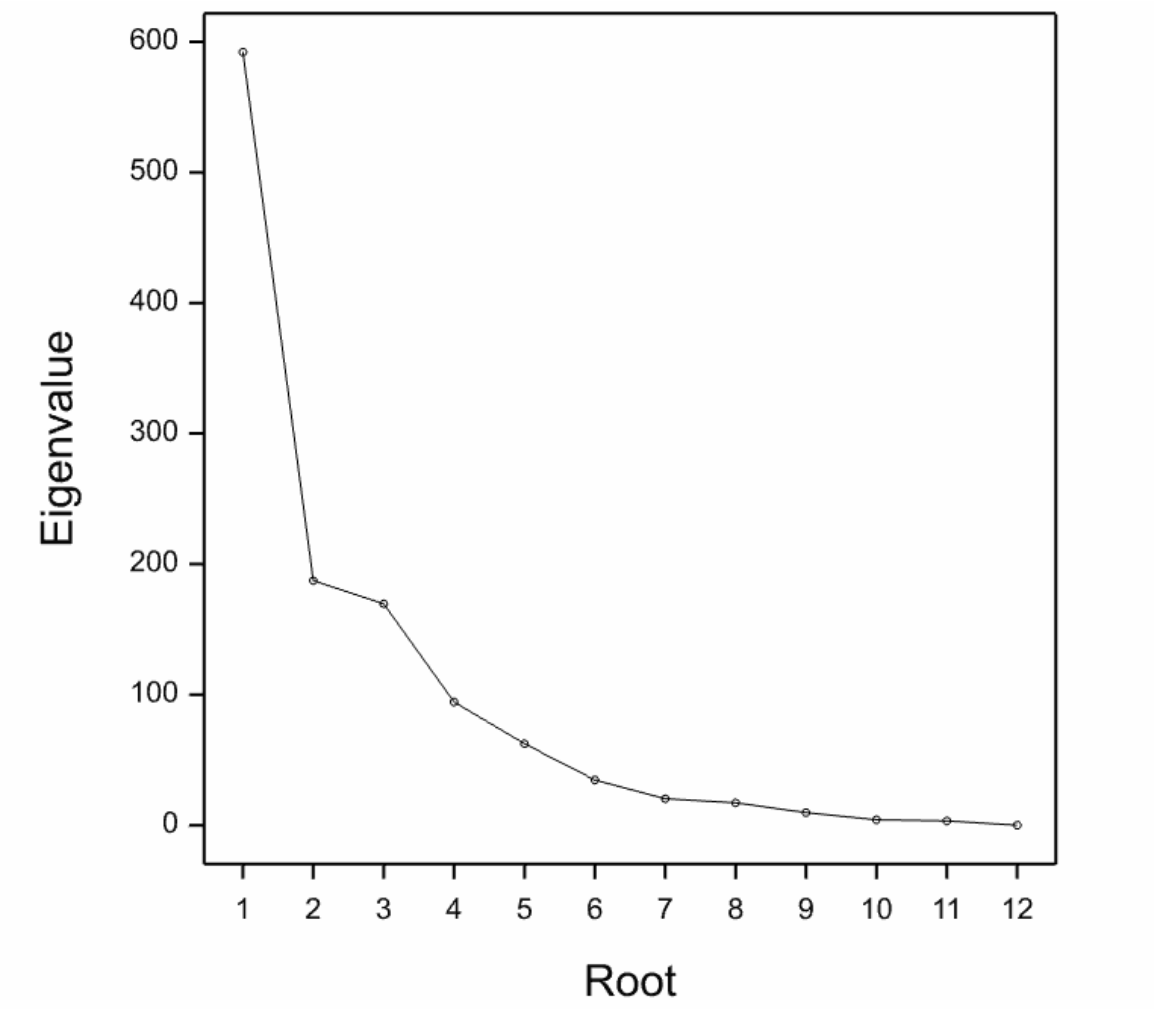

Figure 8.8 Scree plot generated using standardised, normalised transpiration rate (NTR) and fraction of transpirable soil water (FTSW) values from 80 white clover cultivars.

The estimates of correlations calculated for 21 phenotypic traits are visualised in Figure 8.9. There were both positive and negative correlations among traits. The strongest correlation (0.9421) was found among FTSW5 and FTSW7. However, strong correlations existed among FTSW3, FTSW5, FTSW7 and FTSW9. The weakest correlation (-0.565) was among NTR3 and NTR5. Weak correlations were also between FTSW3, FTSW5, FTSW7, FTSW9, FTSW11, FTSW13 and peak number of flowers, seed yield, seed per head and yield. No strong correlations among traits with HCN were found.

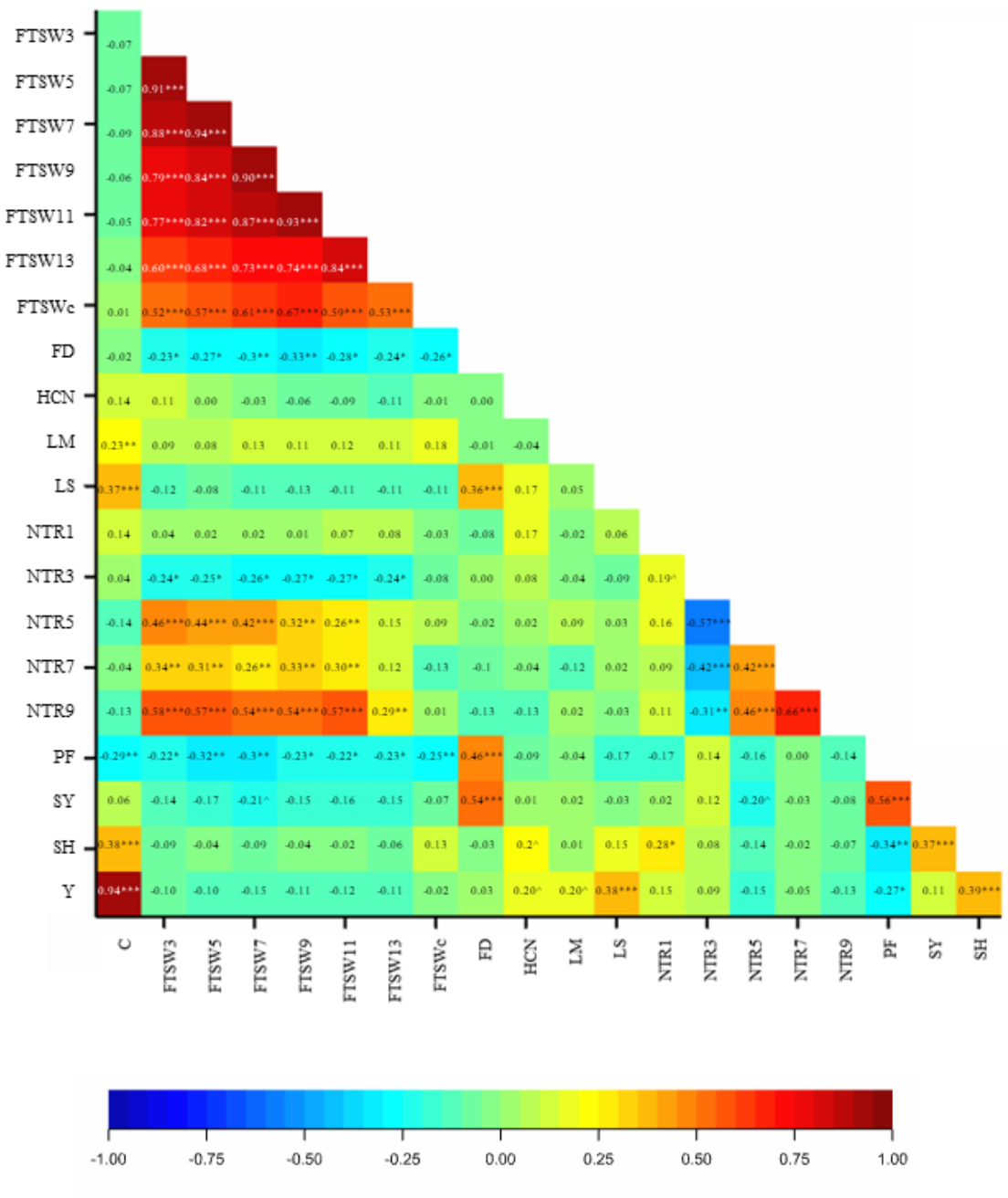

Figure 8.9 Correlation heatmap and correlation r values among the phenotypic traits of content (C), FTSW3, FTSW5, FTSW7, FTSW9, FTSW11, FTSW13, FTSWc, flowering duration (FD), HCN, leaf marking (LM), leaf size (LS), NTR1, NTR3, NTR5, NTR7 NTR9, peak number of flowers (PF), seed yield (SY), seed per head (SH) and dry matter yield (Y) for 80 white clover cultivars. The colours represent the correlation, with red being more positive and blue being more negative.
^*P* < 0·10; **P* < 0·05; ***P* < 0·01; ****P* < 0·001.

8.3.3 Population structure and linkage disequilibrium

To avoid identifying false associations and correct the GWAS analysis for the presence of population structure, the population structure of the panel was analysed. fastStructure inferred that the optimal number of subpopulations (*k*) was 1 when simulated for 1-8 subpopulations. ADMIXTURE confirmed that the lowest cross-validation error was for a subpopulation of *k*=1 (Figure 8.10). A neighbour-joining cladogram showed that the cultivars clustered in 3 groups; 71 cultivars in cluster 1 (black cluster Figure 8.11), 3 cultivars in cluster 2 (yellow cluster Figure 8.11) and 6 cultivars in cluster 3 (pink cluster Figure 8.11).

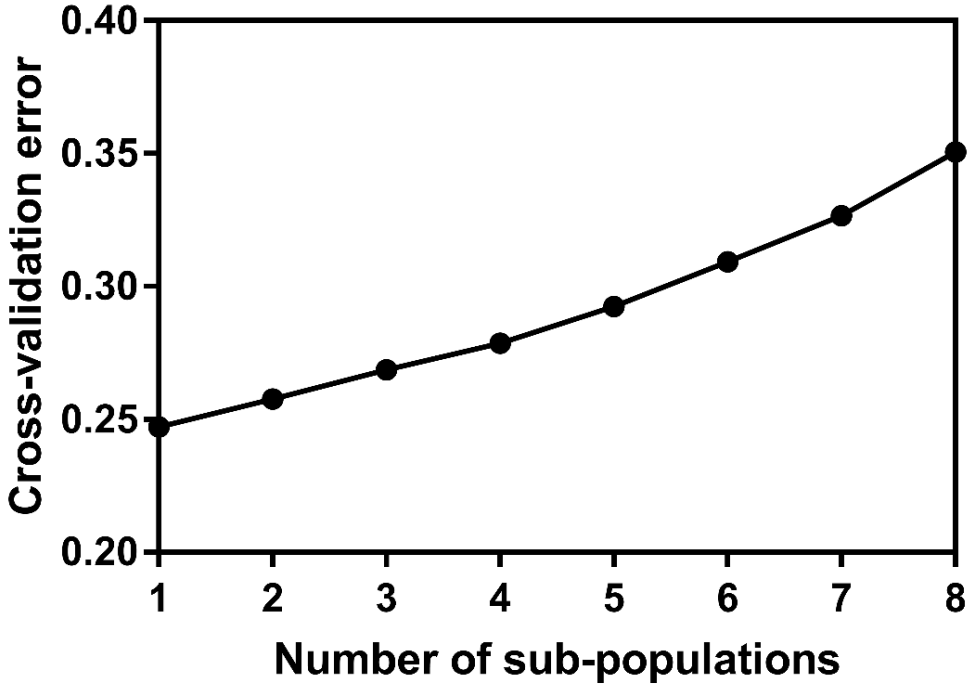

Figure 8.10 The number of subpopulations (1-8) simulated for the 80 white clover cultivars in the diversity panel used in this study. The cross-validation error is on the y-axis, and the number of subpopulations is on the x-axis.

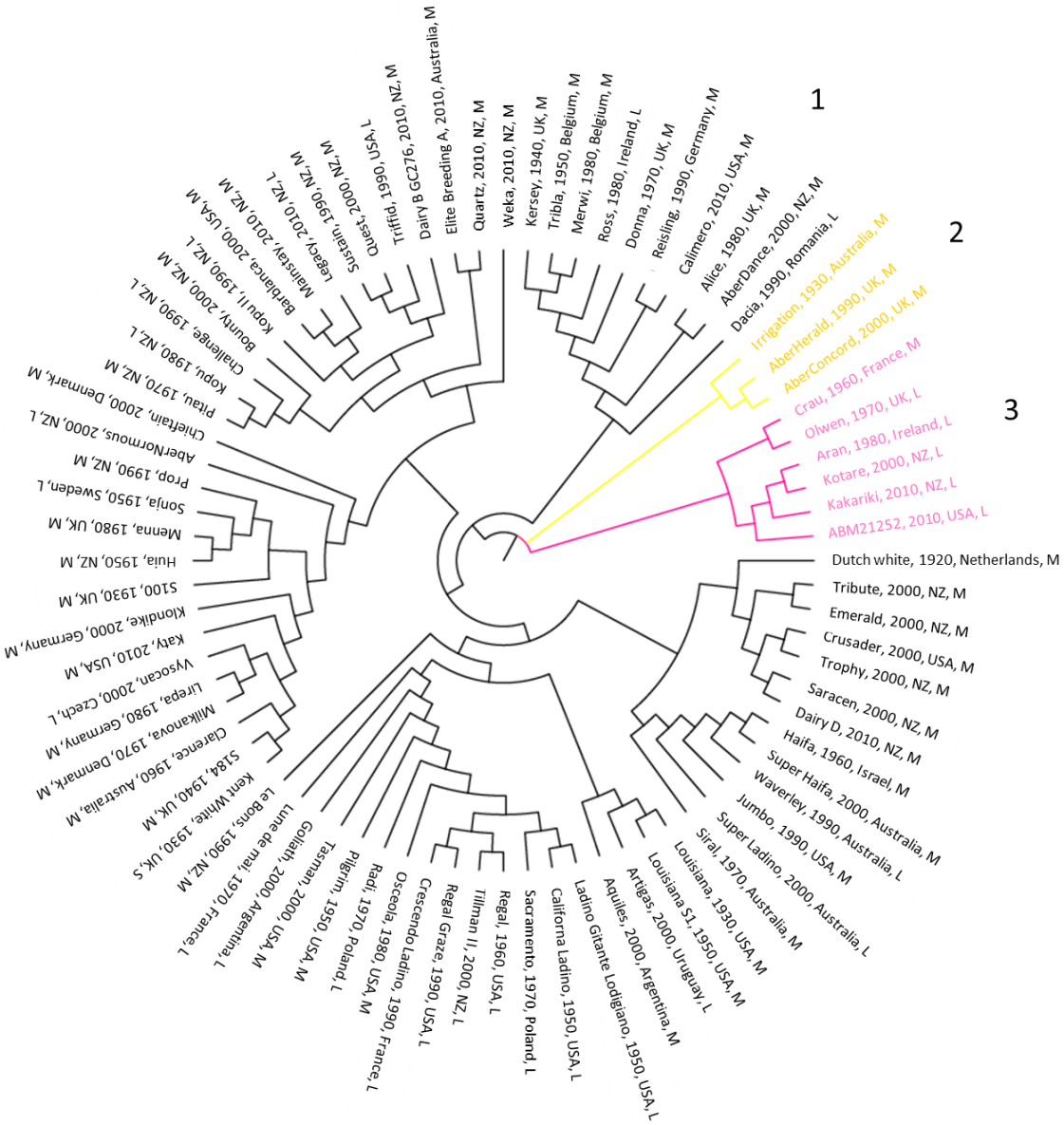

Figure 8.11 The neighbour-joining cladogram of the 80 white clover cultivars in the diversity panel used in this study. Clusters 1, 2 and 3 refers to the different clusters identified.

Cluster 2 (Figure 8.11) contained three cultivars, Irrigation (1930), AberHerald (1990), and AberConcord (2000). The three cultivars were all medium-leaved, and the germplasm associated with these cultivars had been collected from locations that had arid conditions. Both AberConcord and AberHerald were bred by the same company, suggesting similar germplasm was used for breeding both cultivars. However, another cultivar in the panel, AberDance, was bred by the same company as AberConcord and AberHerald but was in a different cluster. AberHerald and AberConcord are bred for high performance under sheep and cattle rotational grazing and are suitable for cutting.

Cluster 3 (Figure 8.11) contained six cultivars; Crau (1960), Olwen (1970), Aran (1980), Kotare (2000), Kakariki (2010) and ABM21252 (2010). All the cultivars were large-leaved apart from Crau. However, Crau is often listed as a large-leaved cultivar (Campbell, et al., 1994, Caradus, 1986,

Caradus, et al., 1997b). The germplasm from Aran, Crau and Olwen originated from France. Specifically, the germplasm from Crau and Olwen originate from southwest France. Aran and Kotare clustered on the same branch. Kotare is considered the natural replacement of Aran. Both cultivars are not tolerant of grazing and are more commonly used for cutting for hay and silage. Crau, Olwen and Aran have moderate-high levels of HCN. Crau, Kakariki and Olwen are similar in morphological traits, particularly for stolon density (Davies, et al., 1992).

The majority of the cultivars (71) clustered in cluster 1 (Figure 8.11). The cluster contained cultivars that have small, medium and large leaf sizes and originate from New Zealand, USA, Israel, Australia, The Netherlands, Uruguay, Argentina, Poland, France, UK, Denmark, Germany, Czech, Sweden, Belgium, Ireland and Romania. One sub-cluster within cluster 1 contained two divergent groups. The first group contained Dutch White, Tribute, Emerald, Crusader, Trophy, Saracen and Dairy D. All these cultivars are medium-leaved and are from the 2000 and 2010 decades. However, Dutch White was separating from the rest of the group. Dutch White was the only cultivar in the panel that was from the 1920s decade. The second cluster contained Siral, Super Ladino, Waverly, Super Haifa, Haifa and Jumbo. All six cultivars had a country of origin with dry, arid conditions. Siral, Super Ladino, Waverly and Super Haifa originate from Australia, and Haifa originates from Israel. Jumbo originates from the USA, but is described as having excellent heat tolerance (Caradus, et al., 1997b).

The cultivars Ladino Gitante Lodigiano, California Ladino, Sacramento, Regal, Tillman II, Regal Graze and Crescendo Ladino, formed a sub-cluster within cluster 1. These cultivars are all listed as large-leaved cultivars. In a similar clustering pattern, S184 and Kent White clustered on the same branch. Kent White was the only small-leaved cultivar in the panel while S184 is often described as small-medium leaved. Leaf size could be a factor influencing the clustering patterns.

The cultivar Le Bons was separated from the major sub-clusters. The parental cultivars of Le Bons include Irrigation, Ross and Pitau (Caradus, et al., 1997b). However, the parental cultivars were not in the same cluster as Le Bons. Caradus, et al. (1997b) states that up to 5% of the leaves could be multifoliate. The multifoliate nature could be separating the cultivar from the other clusters. In a similar pattern, Chieftain was clustering separately in a different sub-cluster. Chieftain was the only recent cultivar from Denmark.

Mainstay, Barblanca and Kopu II clustered in the same group. Barblanca and Kopu II were the parental cultivars of Mainstay (Crush, et al., 2015). Mainstay and Barblanca were clustered on the same branch. As explained in chapter 6, the Mainstay population was diverged by either a Barblanca- or Kopu II-type phenotype. As Mainstay and Barblanca clustered together, it could suggest that the Barblanca phenotype was more desirable for the population when developing the Mainstay populations. Within the same cluster, Kopu II, Barblanca, Mainstay and Legacy clustered together and are cultivars that are described as high yield cultivars (Chapman, et al., 2017, Ford, et al., 2015, Woodfield, et al., 2001b, Woodfield, et al., 2006).

Tribla and Merwi clustered on the same branch in cluster 1 and were the only two cultivars in the panel that originated from Belgium. Meanwhile, Weka clustered separately from the major sub-clusters. Weka was originally bred to replace Sustain, which is in the same larger cluster, although it has outperformed Sustain in yield (Agriseeds, 2009). Weka is high-performing and has high total yield across all seasons and all grazing systems (Caradus, et al., 1996b).

Analysis of LD decay across all genotypes was based on a total of 69,202 SNPs. The average LD decay for the entire genome was estimated at 150bp at r^2=0.2. There was no significant difference (p>0.05) in the rate of LD decay between large and medium leaf sizes.

8.3.4 GWAS of phenotypic traits

Associations were tested between 69,202 SNPs and each trait. The results of GWAS are summarised by Manhattan plots reported for traits in Figure 8.12. The QQ plot results in Figure 8.13 indicated that population stratification was adequately controlled. The majority of the QQ plots do not deviate from the expected chi-squared value, apart from the tail area. The FDR adjusted p-values determined a significant (p<0.05) signal for cyanogenesis (Figure 8.14a and 8.14b). The signal peaked at the SNP 5_45340519, which is located on chromosome 5 at position 45340519. This SNP had a MAF of 0.12. There has been no gene annotation for this SNP. However, when the marker was tested against SNP dosage for significance, it was not significant. No significant marker-trait associations were found for yield, clover content, seed yield, flowering duration, the peak number of flowers, leaf size, seed per head, leaf marking, normalised transpiration rate and the fraction of transpirable soil water.

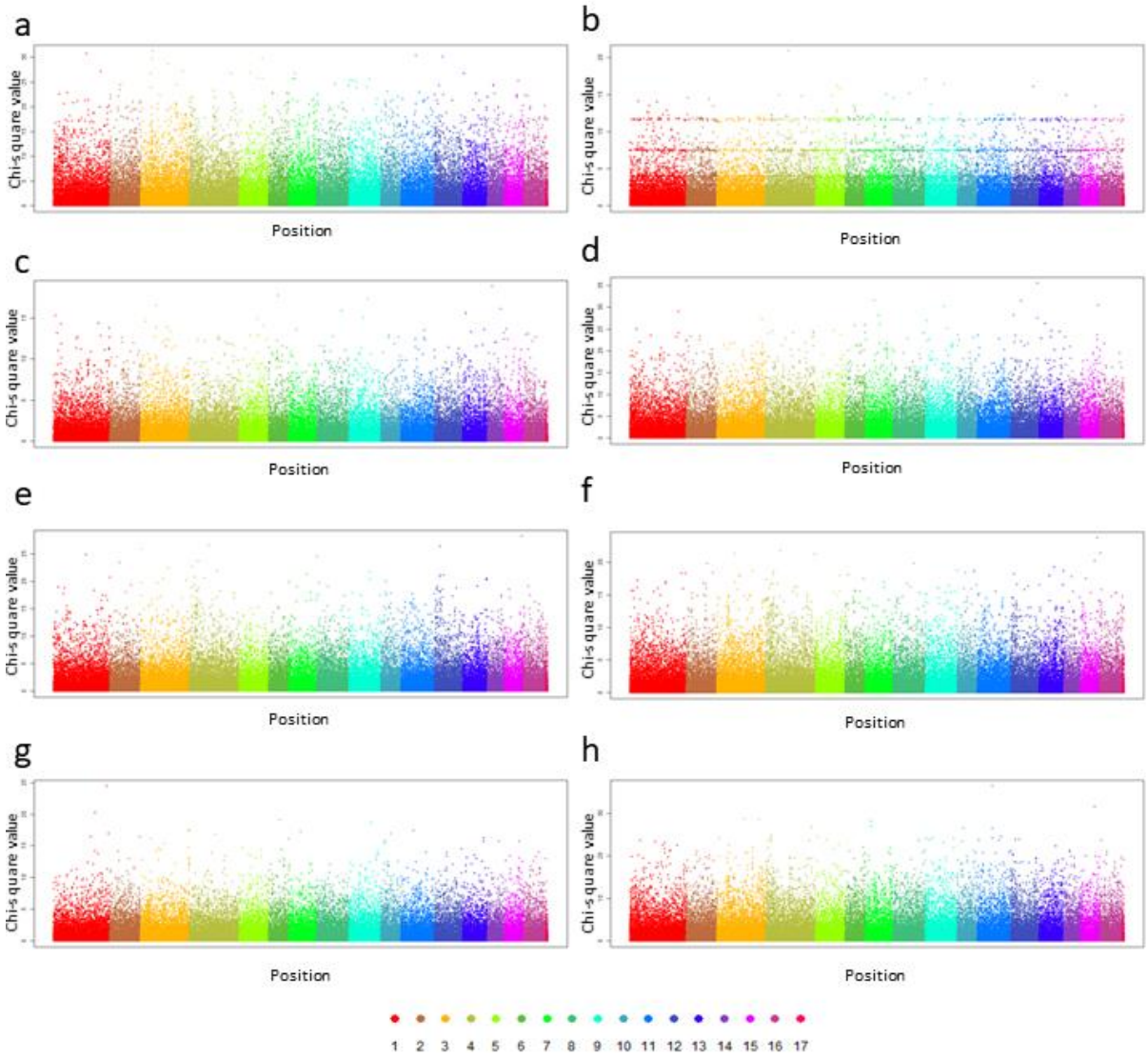

Figure 8.12 Manhattan plots for the phenotypic traits with no significant associations: (a) clover content, (b) flowering duration, (c) leaf marking, (d) leaf size, (e) peak number of flowers, (f) grams of seed per head, (g) seed yield, and (h) yield.

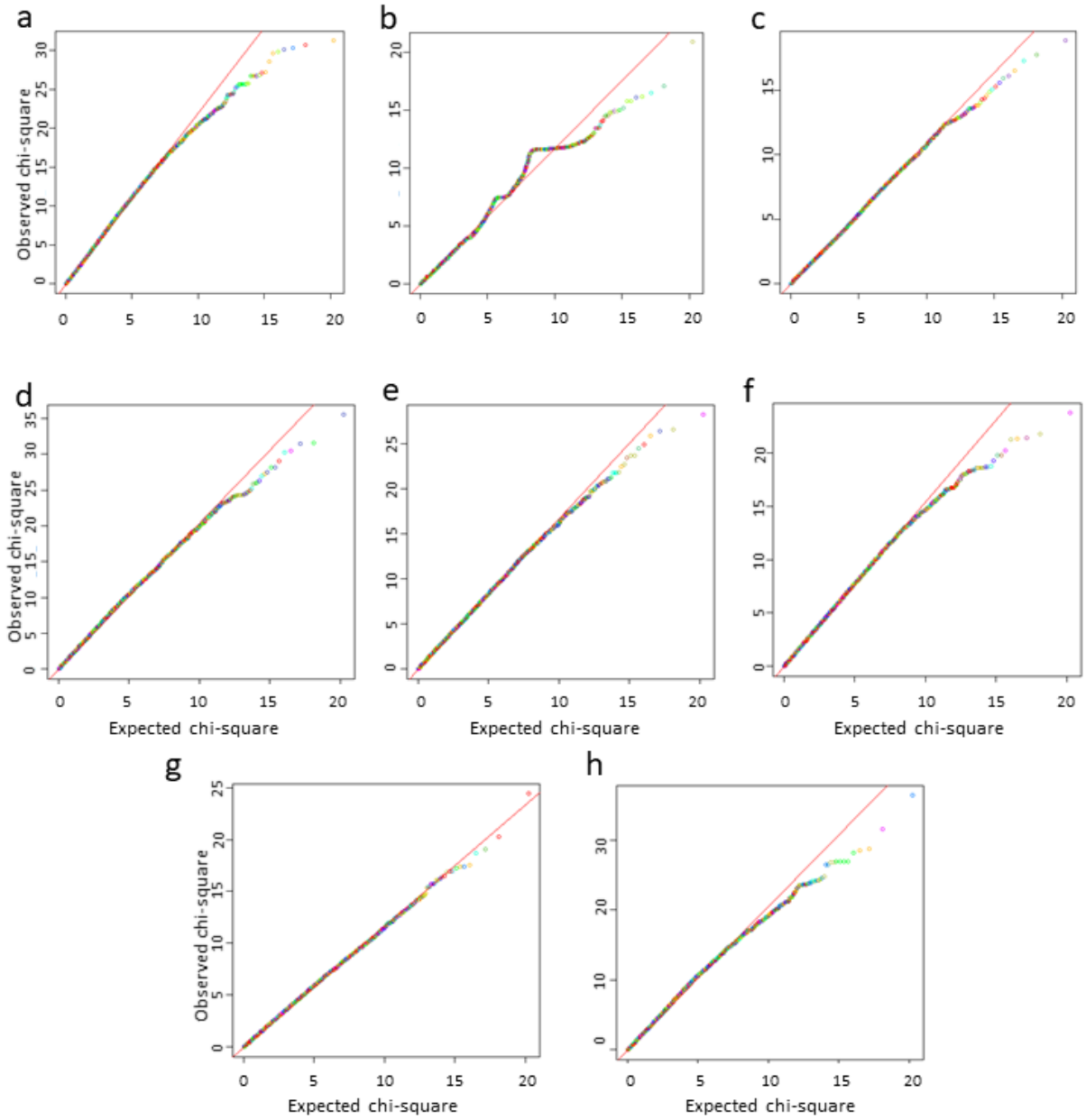

Figure 8.13 Quantile-quantile (QQ) plots for the phenotypic traits with no significant associations: (a) clover content, (b) flowering duration, (c) leaf marking, (d) leaf size, (e) peak number of flowers, (f) grams of seed per head, (g) seed yield, and (h) yield.

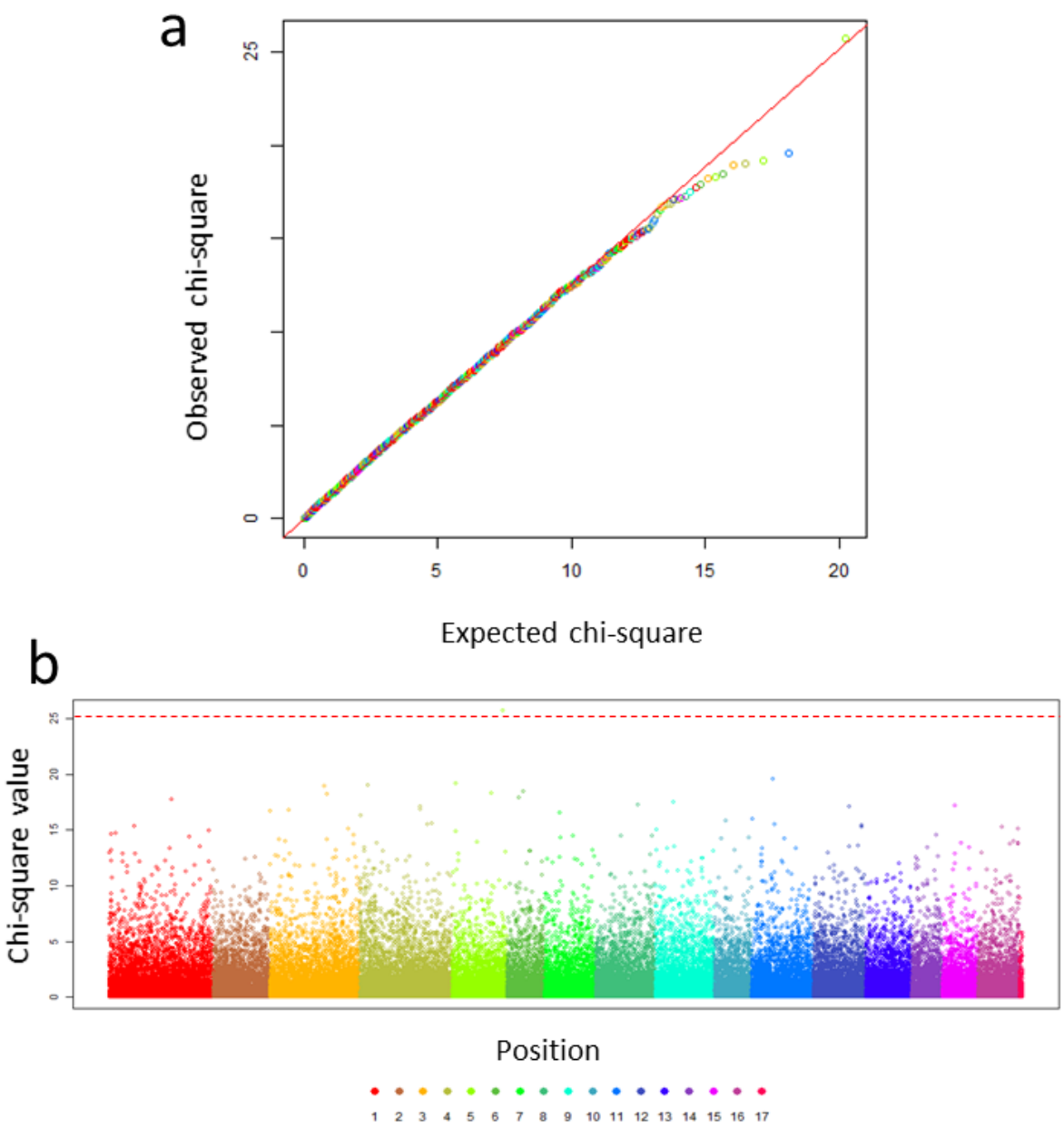

Figure 8.14 Quantile-quantile (QQ) (a) and Manhattan plot (b) for the significant marker-trait association of cyanogenesis. The red dashed line indicates the significance threshold.

8.4 Discussion

8.4.1 Population structure and linkage disequilibrium

The effects of population structure are a frequent concern in GWAS studies. Accounting for population structure is essential in mitigating false-positive associations. Ancestry differences and cryptic relatedness are the two types of relatedness that can produce false positives (Sul, et al., 2018).

White clover is a highly outcrossing species, with a rapid mutation rate, that is dependent on genetic drift and new variation. Although only one subpopulation was identified, Figure 8.11 suggests that if we had more genotypes, there could be three possible clusters. The country of origin of the germplasm

and leaf size is a clustering pattern for cluster 3 in Figure 8.11. All cultivars within the cluster are large-leaved, and cultivars within the cluster are clustering together based on the country of origin of the germplasm. Pedigree analysis of white (Egan, et al., 2019a) and red clover (Egan, et al., 2019b) showed that the accessions clustered together based on country of origin. Furthermore, the influential accessions had a strong influence over the population structure due to the country of origin.

However, the clustering patterns in Figure 8.11 suggest that the between-population variation was low. The low rate of between-population variation is well acknowledged. Population genetics studies in white clover have shown that a large percentage (70-90%) of total genetic variability was explained by within-population variation (Annicchiarico, et al., 2014, Collins, et al., 2002, George, et al., 2006, Gustine, et al., 2001, Olsen, et al., 2007). Collins, et al. (2012) investigated population genetic diversity and structure in white clover. There were no consistent changes in average genetic diversity between populations. The identification of only one subpopulation supports studies that identify low levels of between-population in white clover. Studies have found no significant genetic differentiation in global white clover collections, including natural populations and cultivars (George, et al., 2006, Olsen, et al., 2007).

White clover has only been subject to breeding and selection since the early 1900s. Natural populations of white clover are self-incompatible and heterozygous at many loci, ensuring high levels of variation throughout a population (Burdon, 1980, Phosphate, et al., 1962, Turkington, et al., 1979). The phenotypic plasticity of white clover is widely recognised (Caradus, 1994, Caradus, et al., 1993, Huber, et al., 2008, Hutchings, et al., 1997, Lotscher, et al., 1997, Welham, et al., 2002). The variation present in white clover ensures that populations are adaptable to a wide range of abiotic and biotic conditions, often within a single farm system. The methods of crossing used in white clover breeding programmes, such as polycrosses, are designed to keep a wide genetic base to retain heterozygous genetic combinations (Baker, et al., 1987). The high level of variation present in white clover populations is a challenge for genetics-based studies. New methodologies are needed to overcome the variation so genetics can help to advance white clover populations.

Gage, et al. (2017) assessed the effect of selection on GxE variation and characterised polymorphisms associated with plasticity in maize. They found that regions of the genome that had been selected for in modern breeding programmes explained less variability for yield GxE than unselected regions. They hypothesised that the development of elite populations through breeding could have reduced the GxE of modern cultivars. The genomic positions associated with stability showed fewer associations and enrichment of variants, thought to be due to the control of plasticity by short-range regulatory elements. White clover could benefit from a similar study to determine phenotypic plasticity and quantitatively describe the effect of large, multi-environment trials on white clover populations.

As explained in chapter 6, there are limited publications available on LD patterns and decay rate in white clover. However, a rapid rate of LD decay is common in outcrossing species (Brazauskas, et al.,

2010, Inostroza, et al., 2018, Li, et al., 2014b, Sakiroglu, et al., 2012, Xing, et al., 2007). The rapid rate of LD decay suggests that there is a high rate of recombination, producing high levels of variation. However, the rate of LD decay in the Genetic Gain panel, and also the Mainstay panel in Chapter 6, are more rapid than any reported literature in an outcrossing forage species. There may be an artificial enhancement of LD decay due to approximately 20% of the scaffolds are in the wrong sub-genome.

8.4.2 Genotype-phenotype associations

Overall, the GWAS results confirmed the expected polygenic control of forage traits in white clover. The only SNP to pass the multiple testing threshold was for cyanogenesis. Sufficient phenotypic and genotypic variation was present to allow the signal to pass multiple testing. The highly polymorphic nature of cyanogenesis in white clover was identified in the early 1900s (Armstrong, et al., 1913, Ware, 1925). Cyanogenesis is a simple trait controlled by two genes, and wild populations are often polymorphic at both loci (Daday, 1954a). As cyanogenesis is controlled by two genes, the signal could still be detected in a small sample size. The ecological factors that favour cyanogenic and acyanogenic plants have been widely studied and are one of the most well-documented examples of an adaptive polymorphism in plants.

It is not clearly understood what factors favour acyanogenic white clover types, but there is a higher frequency under cool growing conditions. Cyanogenesis could be deleterious in cooler climates as freezing would cause the cell to rupture and would result in HCN toxicity within the plant (Daday, 1954a, Daday, 1954b). In cooler climates, there are fewer herbivore predators, so there is no advantage for a plant to have that chemical defence (Kakes, 1997). Daday (1965) concluded that cyanogenesis is correlated with increased physiological and morphological fitness in warmer climates. Both loci are associated with polymorphisms for polygenic physiological and morphological traits (Olsen, et al., 2007). The wide range of countries of origin, and associated climates, of the cultivars in the GWAS panel aided in providing enough phenotypic diversity and a subsequent genomic shift in the panel, causing a significant signal to be detected.

All the traits analysed in this chapter, apart from cyanogenesis and leaf size, are described as complex traits. The two possibilities of complex traits are that (i) the trait is controlled by many rare variants, having a large joint effect on the phenotype, or, (ii) many common variants that have a small joint effect on the phenotype. Many traits in outcrossing species are polygenic. No significant signals on the other traits were likely detected because the panel was underpowered by genotypes. When variants are at low frequency or have a small effect size, the considerations of genetic heterogeneity, mutations at two or more loci that produce similar phenotypes (Holland, 2007), and sample size are needed to increase the power of a GWAS to detect significant associations. In complex traits, different environments may influence different variants that control a trait. The multi-location environment of this trial is needed to assess the agronomic performance of the cultivars. But, the genetic heterogeneity will reduce the power to detect the association because the correlation between the phenotype and the

variant is weakened (Korte, et al., 2013). Ultimately, only one GWAS signal was found in this panel due to the small sample size.

Future directions

GWAS continues to be one of the most popular tools of choice for QTL and gene discovery. However, improvements in techniques are allowing more precise and reproducible discoveries. New technologies will provide more detailed phenotypic and genotypic data and could enhance the potential of results. New computational methods are the most prominent advancement of GWAS. The improvements include multi-locus and multi-trait analysis and joint linkage association mapping. Turley, et al. (2018) proposed a method termed Multi-Trait Analysis of GWAS, which can handle multiple traits and rapidly estimates trait-specific effects for individual SNPs. Meta-analyses, often referred to as power analyses, are becoming commonplace in the 'post GWAS era'. Often, datasets that were produced in earlier GWASs are utilised to increase the power to identify causal SNPs. GWAS meta-analyses have two primary objectives: (i) discover new marker-traits associations, and (ii) replicate analyses to confirm marker-trait associations (Gupta, et al., 2019). However, new phenotyping technologies are aiding in dissecting complex traits. High-throughput phenotyping technologies including visible imaging, thermal infrared imaging and 3D imaging, allow increased volume and complexity of data to be collected to understand complex quantitative traits (Du, et al., 2018, Li, et al., 2014a, Yendrek, et al., 2017). The phenotype data collected from the high-throughput phenotyping technologies can be included in GWAS analysis.

8.5 Conclusions

- The population structure of k =1 emphasises that interpopulation variation in white clover is lower than intrapopulation, as seen in chapter 6. Leaf size and country of origin are influencing the clustering patterns within the panel.

- The rapid rate of LD decay is similar to literature with outcrossing forage species. However, the truncated location of 20% of the genome may be inflating the rate of decay.

- A signal was identified for cyanogenesis. The signal was able to be identified in the small panel size as the trait is only controlled by two genes. However, the results from this chapter highlight the complex nature of agronomic traits in white clover. The inability to identify marker-trait associations in the complex traits reinforces the polygenic nature of the traits.

Chapter 9
General conclusions

A key objective of this thesis was to investigate the variation within populations, and between cultivars and germplasm accessions of *Trifolium* species, using both parental and genomic data. Although the *Trifolium* genus is vital to agricultural systems worldwide, there has been minimal published research evaluating the diversity held within accessions (Jahufer, et al., 2013a). The key results of this thesis will have significant implications to pre-breeding decisions for *Trifolium* species. The model showing the key results of this thesis is displayed in Figure 9.1.

An emerging issue with cultivated plant species worldwide is the reduction of variation within populations and the need to monitor and source new variation. Currently, the only published studies using pedigree analysis to estimate and dissect variation in *Trifolium* species are from this thesis (Egan, et al., 2019a, Egan, et al., 2019b, Egan, et al., 2020). Similarly, there is only one published study on a white clover genome-wide associations study (Inostroza, et al., 2018).

The primary aim of this thesis was to evaluate variation within *Trifolium* species populations through both pedigree and genomic analysis. The results from the analysis produced several fundamental findings that will have significant implications for *Trifolium* breeding programmes in the future.

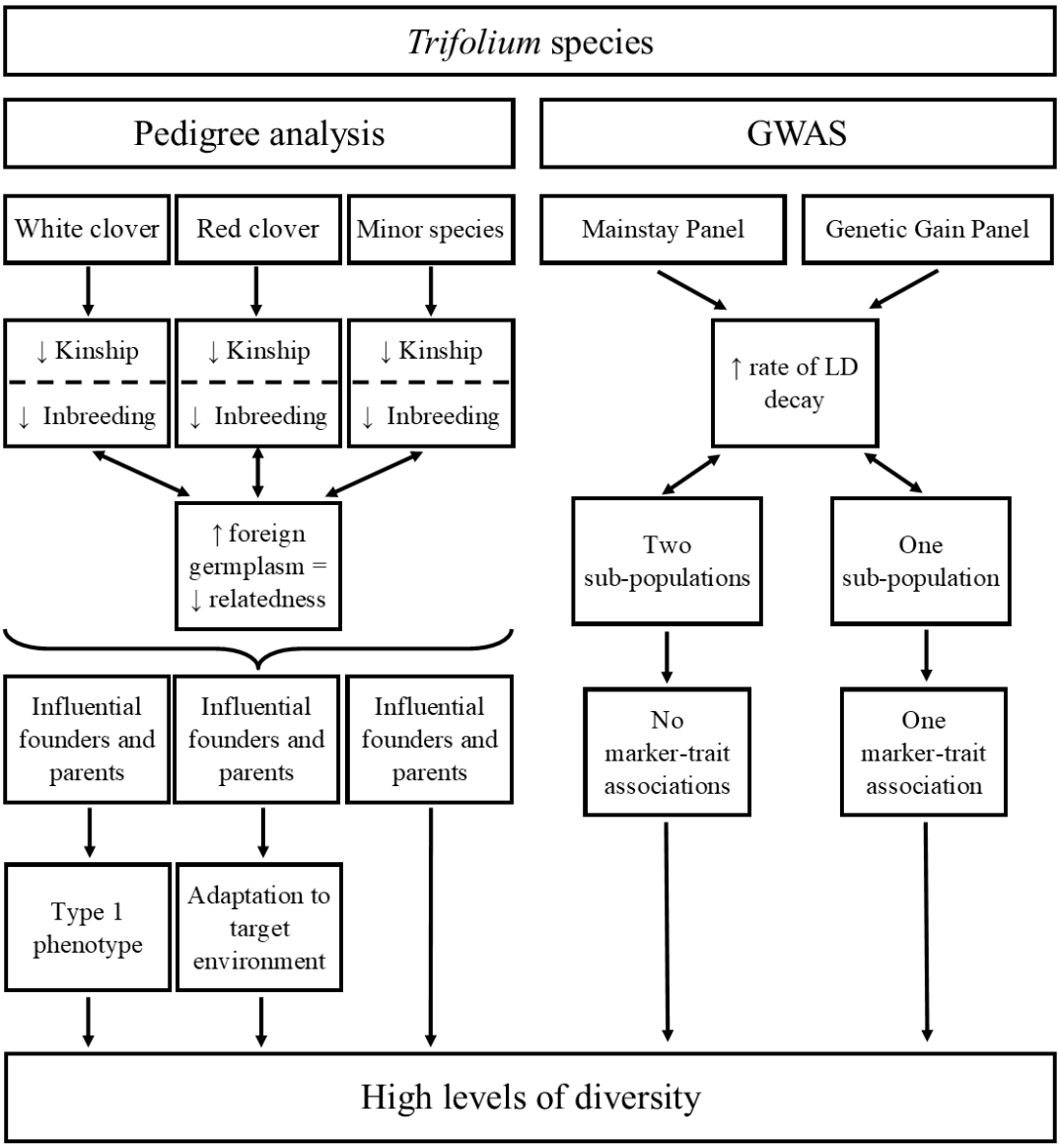

Figure 9.1 The model of the key results of the thesis.

9.1.1 Pedigree analyses

Pedigree analysis is a cost-effective, straightforward and valuable method to monitor variation in species. The creation of pedigree maps allows breeders to visualise the breeding that has occurred and the resulting patterns of breeding decisions. Pedigree maps are a method to detect breeding patterns in programmes and to inform decisions about the next selection cycle (Shaw, et al., 2014, Shaw, et al., 2016). The development of the pedigree maps for *Trifolium* species showed breeding patterns that have occurred since the early 1900s. Surprisingly, no obvious bottlenecks were identified in any species. The pedigree map identified accessions that contributed large numbers of progeny, and they will be monitored for diversity in the future. The interactive nature of the pedigree maps allowed nodes of interest to be isolated and accessions to be traced back to founding accessions.

The calculation of inbreeding and relatedness coefficients showed that there are high levels of diversity among all species. The low levels of inbreeding and relatedness through the *Trifolium*

species is encouraging as it is detrimental for *Trifolium* species to cross with closely related accessions. However, among the inbred plants, there were accessions which had high levels of inbreeding >10%. As high inbreeding levels are unfavourable to *Trifolium* species, these accessions will be monitored in future breeding programmes. Although, the high levels of inbreeding could indicate that the accessions have been involved in a lot of breeding activity, signifying that the accessions could have desirable phenotypic traits.

Important accessions that influenced the population structure were identified by the derivation of kinship coefficients and subsequently identified influencing founders. The breeding pattern and history of the influential accessions could be traced back to significant founders. Over 70% of the founders identified in white clover could be attributed back to the Type 1 phenotype. The Type 1 phenotype is still desirable to modern cultivars and has influenced the population structure of the white clover accessions in the MFGC. In red clover, the influence of the important accessions could be attributed to the adaptation for the target environment in New Zealand. In the minor *Trifolium* species, some of the influence came from introductions from foreign countries. For some of the species, such as *T. hybridum,* which are often hybridised with white clover, accessions that originated from drought-prone areas of New Zealand, such as the Mackenzie country influenced the accessions. In *T. ambiguum*, the Australian cultivar 'Monaro' had a strong influence over the whole population structure. The productive nature of Monaro and the observed drought tolerance could be a reason as to why it is influencing the population.

Although the MFGC has been successful in maintaining *Trifolium* collections with a broad base of variation, the level of variation must be maintained to support future breeding programmes. Germplasm collection trips will remain vital in introducing new variation into the collections and providing novel sources of variation to address new breeding targets. The identification of geographic origins that have influenced the population structure of *Trifolium* species could inform the MFGC on future germplasm collection trips, sampling strategies and the uniqueness of germplasm in relation to other global collections.

The ultimate aim of the results from the pedigree analyses was to provide avenues to improve the rate of genetic gain and environmental adaptation in *Trifolium* species. A low rate of gain is widely acknowledged in *Trifolium* species (Hoyos-Villegas, et al., 2019, Riday, 2010, Tucak, et al., 2013, Woodfield, 1999, Woodfield, et al., 1994) and continual improvements of breeding methods and tools are needed for future studies. The information obtained from these studies will inform and could improve the efficiency of crosses, and design new breeding strategies utilising variation both within and among clusters. The information derived from the pedigree analysis will help improve domestic and global germplasm exchange efforts and utilisation at the MFGC. The characterisation of the germplasm will enable breeders to have increased efficiency for crosses and could ultimately shorten the length of the breeding programme.

9.1.2 Genomic analyses

In recent decades, there has been a substantial increase in the amount of GWASs in animal and plant species (Visscher, et al., 2012, Visscher, et al., 2017). GWASs has become the first step in the genetic analysis and dissection of complex traits. The identification of a specific causal variant is possible with sufficient genotypes and power. A crucial aspect in attempting to increase the rate of genetic gain of white clover is the utilisation of genomic tools such as GWAS. Considering this, the lack of published studies on GWASs in white clover was surprising. Although there have been several studies identifying QTL for agronomic traits (Barrett, et al., 2005, Barrett, et al., 2009, Faville, et al., 2012, Isobe, et al., 2009, Williams, et al., 2007), there is only one study identifying marker-trait associations (Inostroza, et al., 2018). However, the results from this thesis showed the high levels of phenotypic and genotypic plasticity of white clover and could explain the lack of marker-trait associations.

Two panels were analysed for important agronomic traits. The Mainstay panel was a single breeding population, and the Genetic Gain panel was a diversity panel with 80 commercial cultivars. The population structure of the Mainstay panel showed two subpopulations and would be likely attributed to the two parental cultivars, whereas the Genetic Gain panel showed no population structure. The minimal or lack of population structure in both panels reiterates the high levels of diversity that are present in white clover due to the mating system. The highly outcrossing mating systems produce populations that are highly variable, both within and between populations. The rapid rate of LD decay in both panels supports the high levels of variation. Rapid decay of LD is common in outcrossing forage species. White clover is a species that relies on new variation and genetic drift throughout a population. The outcrossing breeding system ensures that there is a broad genetic base and high recombination rate, resulting in a rapid rate of LD decay.

The results confirmed that important agronomic traits in white clover have complex genetics and could explain the lack of published studies for GWASs in white clover. One marker-trait association was identified for the simple trait of cyanogenesis, showing that the SNP pipeline was successful. However, the complex agronomic traits are often of high importance to white clover breeding programmes. Increased numbers of genotypes are needed to increase the power of the study.

9.2 Limitations

The limitations of this research are summarised below:

- The major constraint for the pedigree map construction and pedigree analysis was the quality of the records maintained. The pedigree records started in the early 1900s for some *Trifolium* species and the descriptive techniques of how breeders record the crosses have changed over time. The lack of recording of crosses meant that some germplasm could not be included in the analyses.

- A limitation in the analysis of the pedigree data was that the pedigree analysis software could only infer the data to have two parents. Often in forage breeding, techniques are used where there are more than two parents, i.e. polycrosses. Therefore, accessions with more than two parents had to be excluded from the analyses (Figures 3.1 and 4.1).

- The Mainstay panel was from a single breeding population. Usually, GWASs are performed on diversity panels to ensure there is sufficient genotypic and phenotypic diversity. The lack of phenotypic and genotypic variation within the panel confirmed that there were no significant marker-trait associations.

- Phenotypic variation is needed in a GWAS panel to signal a genomic shift. The scoring scale of phenotypes is significant when harbouring variation. If the scale on which phenotypic traits are scored do not have a large range, there may not be enough variation to reflect a significant phenotypic and genotypic shift. The stolon density measurements in the Mainstay panel were measured on a 1-5 scale, which may not have been diverse enough for this study. Future studies will utilise a scale of 1-9.

- The main limitation in the Genetic Gain GWAS panel was the number of genotypes. Many of the traits analysed were complex and polygenic, and the variants could be at a low frequency or have a small effect size. An increased sample size would have increased the power of the study to detect associations.

9.3 Future applications

Based on this work, future research avenues are as follows:

- Pedigree analysis on ryegrass accessions held at the MFGC. Ryegrass and white clover grow in companionship in a pastoral system. Like the *Trifolium* species, there is limited published research on evaluating the variation present in ryegrass collections (Casler, 1995). There is speculation that ryegrass has been through significant breeding bottlenecks and this is of concern to breeders.

- The construction of complex pedigree maps and performing pedigree analysis in populations can provide genetic studies with substantial analytical power (Fradgley, et al., 2019). Pedigree-based approaches for mapping rare variants have advantages when compared to population-based methods. Rare variants are often shared by many relatives in a pedigree and will, therefore, be in abundance in the study, improving the power to identify and confirm causal variants. Studies could also be designed to focus on pedigree lineages that are elite or highly variable in certain traits (Vinson, et al., 2013). Sequencing key influential or important accessions in a pedigree could reveal loci controlling traits of interest (Ma, et al., 2019).

- The integration of pedigree analysis into GWAS studies is a method termed selection mapping. Selection mapping refers to methodologies that identify alleles, loci and epistatic interactions in populations that have been subject to cycles of selection (Wisser, et al., 2008). A selection mapping technique is often used when establishing the genetic diversity and relationships in a panel and to identify genomic regions associated with traits (Gutierrez-Gil, et al., 2014, Liu, et al., 2016, Shugart, et al., 2012, Soleimani, et al., 2002, Sud, et al., 2005, Zhou, et al., 2014).

- Pedigree information has been used to track the inheritance of markers (Matsumoto, et al., 2017). New marker-assisted breeding programmes have been developed that have come from gene and QTL discovery on breeding germplasm based on pedigree analysis termed pedigree-based QTL mapping (Liu, et al., 2016). The use of marker information on pedigrees increases the resolution of QTL discovery and helps to assemble germplasm panels to use in studies (Laurens, et al., 2018).

- The improvement of software packages, especially for pedigree analysis, will allow for more accessible communication for germplasm exchange and may lead to an increase in utilisation and knowledge of the germplasm.

References

Abbasi, M. 2008. Genetic diversity of Persian clover (*Trifolium resupinatum*) gene pools in National Plant Gene Bank of Iran.

Abberton, M., T. Michaelson-Yeates, A. Marshall, K. Holdbrook-Smith and I. Rhodes. 1998. Morphological characteristics of hybrids between white clover, *Trifolium repens* L., and Caucasian clover, *Trifolium ambiguum* M. Bieb. Plant Breeding 117: 494-496. doi:10.1111/j.1439-0523.1998.tb01981.x.

Abberton, M.T. 2007. Interspecific hybridization in the genus *Trifolium*. Plant Breeding 126: 337-342. doi:10.1111/j.1439-0523.2007.01374.x.

Abberton, M.T., J.H. MacDuff, S. Vagg, A.H. Marshall and T. Michaelson-Yeates. 2000. Nitrogen fixation in hybrids of white clover (*Trifolium repens* L.) and Caucasian clover (*Trifolium ambiguum* M. Bieb). Journal of Agronomy and Crop Science 185: 241-247. doi:10.1046/j.1439-037x.2000.00438.x.

Abberton, M.T. and A.H. Marshall. 2005. Progress in breeding perennial clovers for temperate agriculture. The Journal of Agricultural Science 143: 117-135. doi:10.1017/S0021859605005101.

Abberton, M.T. and I. Thomas. 2011. Genetic resources in *Trifolium* and their utilization in plant breeding. Plant Genetic Resources 9: 38-44. doi:10.1017/S1479262110000341.

Abdel-Haleem, H., T.E. Carter, L.C. Purcell, C.A. King, L.L. Ries, P. Chen, et al. 2012. Mapping of quantitative trait loci for canopy-wilting trait in soybean (*Glycine max* L. Merr). Theoretical and Applied Genetics 125: 837-846. doi:10.1007/s00122-012-1876-9.

Abdi, A.I., P.G. Nichols, P. Kaur, B.J. Wintle and W. Erskine. 2020. Morphological diversity within a core collection of subterranean clover (*Trifolium subterraneum* L.): Lessons in pasture adaptation from the wild. PloS One 15: e0223699. doi:10.1371/journal.pone.0223699.

Acosta-Gallegos, J.A., J.D. Kelly and P. Gepts. 2007. Prebreeding in common bean and use of genetic diversity from wild germplasm. Crop Science 47: S-44-S-59. doi:10.2135/cropsci2007.04.0008IPBS.

Acquaah, G. 2012. Principles of Plant Genetics and Breeding. 2nd ed. John Wiley & Sons.

Aerts, R.J., T.N. Barry and W.C. McNabb. 1999. Polyphenols and agriculture: beneficial effects of proanthocyanidins in forages. Agriculture, Ecosystems & Environment 75: 1-12. doi:10.1016/S0167-8809(99)00062-6.

Agriseeds. 2009. Weka white clover. Agriseeds Product Sheet.

Aguilera, P.M., M.E. Sartor, F. Galdeano, F. Espinoza and C.L. Quarin. 2011. Interspecific tetraploid hybrids between two forage grass species: sexual *Paspalum plicatulum* and apomictic *P. guenoarum*. Crop Science 51: 1544-1550. doi:10.2135/cropsci2010.10.0610.

Alexander, D.H., J. Novembre and K. Lange. 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome research 19: 1655-1664. doi:10.1101/gr.094052.109.

Allard, R.W., P.D.T. Alvim, A. Ashri, J.H. Barton, F.H. Buttel, T.-T. Chang, et al. 1991. The US National Plant Germplasm System. Managing Global Genetic Resources. National Academies Press, Washington DC.

Allard, R.W., P.d.T. Alvim, J.H. Barton, F.H. Buttel, T.-T. Chang, P.R. Day, et al. 1993a. Documentation of genetic resources. In: N. R. Council, editor Managing global genetic resources: agricultural crop issues and policies. National Academies Press. p. 205-218.

Allard, R.W., P.d.T. Alvim, J.H. Barton, F.H. Buttel, T.-T. Chang, P.R. Day, et al. 1993b. The science of managing genetic resources. In: N. R. Council, editor Managing global genetic resources: agricultural crop issues and policies. National Academies Press. p. 153-172.

Anderson, C.B., B.K. Franzmayr, S.W. Hong, A.C. Larking, T.C. van Stijn, R. Tan, et al. 2018. Protocol: A versatile, inexpensive, high-throughput plant genomic DNA extraction method suitable for genotyping-by-sequencing. Plant Methods 14: 1-10. doi:10.1186/s13007-018-0336-1.

Anderson, M. and N. Taylor. 1974. Effect of temperature on intra-and interspecific crosses of diploid and tetraploid red clover, *Trifolium pratense* L. Theoretical and Applied Genetics 44: 73-76. doi:10.1007/BF00277956.

Annicchiarico, P., B. Barrett, E.C. Brummer, B. Julier and A.H. Marshall. 2015. Achievements and challenges in improving temperate perennial forage legumes. Critical Reviews in Plant Sciences 34: 327-380. doi:10.1080/07352689.2014.898462.

Annicchiarico, P. and M. Carelli. 2014. Origin of Ladino white clover as inferred from patterns of molecular and morphophysiological diversity. Crop Science 54: 2696-2706. doi:10.2135/cropsci2014.04.0308.

Annicchiarico, P., L. Pecetti, H. Bouzerzour, R. Kallida, A. Khedim, C. Porqueddu, et al. 2011. Adaptation of contrasting cocksfoot plant types to agricultural environments across the Mediterranean basin. Environmental and Experimental Botany 74: 82-89. doi:10.1016/j.envexpbot.2011.05.002.

Annicchiarico, P. and E. Piano. 2004. Indirect selection for root development of white clover and implications for drought tolerance. Journal of Agronomy and Crop Science 190: 28-34. doi:10.1046/j.0931-2250.2003.00070.x.

Ansari, H.A., N.W. Ellison, A.G. Griffiths and W.M. Williams. 2004. A lineage-specific centromeric satellite sequence in the genus *Trifolium*. Chromosome Research 12: 357-367. doi:10.1023/B:CHRO.0000034099.19570.b7.

Arcioni, S. and D. Mariotti. 1983. Selfing and interspecific hybridization in *Lolium perenne* L. and *Lolium multiflorum* Lam. Evaluated by phosphoglucoisomerase as isozyme marker. Euphytica 32: 33-40. doi:10.1007/BF00036861.

Armstead, I.P., A. Bollard, W. Morgan, J. Harper, I. King, R. Jones, et al. 2001. Genetic and physical analysis of a single *Festuca pratensis* chromosome segment substitution in *Lolium perenne*. Chromosoma 110: 52-57. doi:10.1007/s004120000122.

Armstrong, H.E., E.F. Armstrong and E. Horton. 1913. Herbage Studies. II.–Variation in *Lotus corniculatus* and T*rifolium repens* (Cyanophoric plants). Proceedings of the Royal Society of London. Series B, Biological Sciences 86: 262-269.

Arojju, S.K., S. Barth, D. Milbourne, P. Conaghan, J. Velmurugan, T.R. Hodkinson, et al. 2016. Markers associated with heading and aftermath heading in perennial ryegrass full-sib families. BMC Plant Biology 16: 160. doi:10.1186/s12870-016-0844-y.

Arseniuk, E. 1989. Effect of the polyploidization red clover (Trifolium pranese L.) on winter hardi ness and resistance to some diseases. Hodowla ros lin, Aklimatizacja Nasienitstwo 33: 1-30.

Ashraf, B.H., S. Byrne, D. Fé, A. Czaban, T. Asp, M.G. Pedersen, et al. 2016. Estimating genomic heritabilities at the level of family-pool samples of perennial ryegrass using genotyping-by-sequencing. Theoretical and Applied Genetics 129: 45-52. doi:10.1007/s00122-015-2607-9.

Asimit, J. and E. Zeggini. 2010. Rare variant association analysis methods for complex traits. Annual Review of Genetics 44: 293-308. doi:10.1146/annurev-genet-102209-163421.

Atwell, S., Y.S. Huang, B.J. Vilhjálmsson, G. Willems, M. Horton, Y. Li, et al. 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature 465: 627. doi:10.1038/nature08800.

Atwood, S.S. 1945. Behaviour of self-compatibility factor and its relation to breeding methods in *Trifolium repens*. Journal of the American Society of Agronomy 37: 991-1004.

Aulchenko, Y.S., M.V. Struchalin and C.M. van Duijn. 2010. ProbABEL package for genome-wide association analysis of imputed data. BMC Bioinformatics 11: 134. doi:10.1186/1471-2105-11-134.

Ayres, J., J. Caradus, R.D. Murison, L. Lane and D. Woodfield. 2007. Grasslands Trophy - a new white clover (*Trifolium repens* L.) cultivar with tolerance of summer moisture stress. Australian Journal of Experimental Agriculture 47: 110-115. doi:10.1071/EA04029.

Azhar, M.T. and A. Rehman. 2018. Overview on effects of water stress on cotton plants and productivity. Biochemical, Physiological and Molecular Avenues for Combating Abiotic Stress Tolerance in Plants: 297-316. doi:10.1016/B978-0-12-813066-7.00016-4.

Aziz, N., N. Paiva, G. May and R. Dixon. 2005. Profiling the transcriptome of alfalfa glandular trichomes. Planta 221: 28-38. doi:10.1007/s00425-004-1424-1.

Bagherzadi, L., T.R. Sinclair, M. Zwieniecki, F. Secchi, W. Hoffmann, T.E. Carter, et al. 2017. Assessing water-related plant traits to explain slow-wilting in soybean PI 471938. Journal of Crop Improvement 31: 400-417. doi:10.1080/15427528.2017.1309609.

Baker, M.J. and W.M. Williams. 1987. Genetics and breeding. White clover. CAB international. p. 343-420.

Barbour, M., J. Caradus, D. Woodfield and W. Silvester. 1996. Water stress and water use efficiency of ten white clover cultivars. Special Publication - Agronomy Society of New Zealand: 159-162.
Barenbrug. 2013. Barblanca White Clover. Forage Tech Sheet.
Barnett, O. and P. Gibson. 1975. Identification and prevalence of white clover viruses and the resistance of *Trifolium* species to these viruses. Crop Science 15: 32-37. doi:10.2135/cropsci1975.0011183X001500010010x.
Barrett, B., I. Baird and D. Woodfield. 2005. A QTL analysis of white clover seed production. Crop Science 45: 1844-1850. doi:10.2135/cropsci2004.0679.
Barrett, B., I. Baird and D. Woodfield. 2009. White clover seed yield: a case study in marker-assisted selection. Molecular Breeding of Forage and Turf: 241-250. doi:10.1007/978-0-387-79144-9_22.
Barrett, B., M. Faville, A. Sartie, D. Hume, Z. Jahufer, M. Hickey, et al. 2006. Forage improvement via marker-assisted selection. Advances in Pasture Plant Breeding. Grassland Research and Practice Series 12: 11-15.
Barrett, B., A. Griffiths, C. Mercer, N. Ellison, M. Faville, S. Easton, et al. 2001. Marker-assisted selection to accelerate forage improvement. New Zealand Grassland Association: 241-246.
Barrett, B., A. Griffiths, M. Schreiber, N. Ellison, C. Mercer, J. Bouton, et al. 2004. A microsatellite map of white clover. Theoretical and Applied Genetics 109: 596-608. doi:10.1007/s00122-004-1658-0.
Basigalup, D., D. Barnes and R. Stucker. 1995. Development of a core collection for perennial *Medicago* plant introductions. Crop Science 35: 1163-1168. doi:10.2135/cropsci1995.0011183X003500040042x.
Belaygue, C., J. Wery, A. Cowan and F. Tardieu. 1996. Contribution of leaf expansion, rate of leaf appearance, and stolon branching to growth of plant leaf area under water deficit in white clover. Crop Science 36: 1240-1246. doi:10.2135/cropsci1996.0011183X003600050028x.
Bell, A.E. 1977. Heritability in Retrospect. The Journal of Heredity 68: 297-300. doi:10.1093/oxfordjournals.jhered.a108840.
Bellaloui, N., A.M. Gillen, A. Mengistu, H. Kebede, D.K. Fisher, J.R. Smith, et al. 2013. Responses of nitrogen metabolism and seed nutrition to drought stress in soybean genotypes differing in slow-wilting phenotype. Frontiers in Plant Science 4: 498. doi:10.3389/fpls.2013.00498.
Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: series B (Methodological) 57: 289-300.
Bennett, M.D. and I.J. Leitch. 2011. Nuclear DNA amounts in angiosperms: targets, trends and tomorrow. Annals of Botany 107: 467-590. doi:10.1093/aob/mcq258.
Bennetts, H., E. Uuderwood and F. Shier. 1946. A specific breeding problem of sheep on subterranean clover pastures in Western Australia. Australian Veterinary Journal 22: 2-12. doi:10.1111/j.1751-0813.1946.tb15473.x.
Bermejo, R., J.J. Irigoyen and J.M. Santamaría. 2006. Short-term drought response of two white clover clones, sensitive and tolerant to O_3. Physiologia Plantarum 127: 658-669. doi:10.1111/j.1399-3054.2006.00695.x.
Bernardo, R.N. 2010. Breeding for quantitative traits in plants. 2nd ed. Stemma Press, Woodbury, Minn.
Beuselinck, P., J. Bouton, W. Lamp, A. Matches, M. McCaslin, C. Nelson, et al. 1994. Improving legume persistence in forage crop systems. Journal of Production Agriculture 7: 311-322.
Biazzi, E., N. Nazzicari, L. Pecetti, E.C. Brummer, A. Palmonari, A. Tava, et al. 2017. Genome-wide association mapping and genomic selection for alfalfa (*Medicago sativa*) forage quality traits. PloS One 12. doi:10.1371/journal.pone.0169234.
Bink, M., P. Uimari, M. Sillanpää, L. Janss and R. Jansen. 2002. Multiple QTL mapping in related plant populations via a pedigree-analysis approach. Theoretical and Applied Genetics 104: 751-762. doi:10.1007/s00122-001-0796-x.
Bishop, J. and M. Korn. 1969. Natural selection and cyanogenesis in white clover, *Trifolium repens*. Heredity 24: 423-430.
Bissuel-Belaygue, C., A.A. Cowan, A.H. Marshall and J. Wery. 2002a. Reproductive development of white clover (*Trifolium repens* L.) is not impaired by a moderate water deficit that reduces

vegetative growth. I. Inflorescence, floret and ovule production. Crop Science 42: 406-414. doi:10.2135/cropsci2002.0414.
Bissuel-Belaygue, C., A.A. Cowan, A.H. Marshall and J. Wery. 2002b. Reproductive development of white clover (*Trifolium repens* L.) is not impaired by a moderate water deficit that reduces vegetative growth. II. Fertilization efficiency and seed set. Crop Science 42: 414-422. doi:10.2135/cropsci2002.0414.
Blows, M.W. and A.A. Hoffmann. 2005. A reassessment of genetic limits to evolutionary change. Ecology 86: 1371-1384. doi:10.1890/04-1209.
Blum, A. 2011. Drought resistance–is it really a complex trait? Functional Plant Biology 38: 753-757. doi:10.1071/FP11101.
Bolle, H.-J. 2003. Climate, climate variability, and impacts in the Mediterranean area: an overview. Mediterranean Climate. Springer. p. 5-86.
Bouton, J. 2012. Breeding lucerne for persistence. Crop and Pasture Science 63: 95-106. doi:10.1071/CP12009.
Bouton, J.H. 2010. Future developments and uses. Fodder Crops and Amenity Grasses. Springer. p. 201-209.
Bouton, J.H., D.R. Woodfield, J.R. Caradus and D.T. Wood. 2005. Registration of 'Durana' white clover. Crop Science 45: 797-797. doi:10.2135/cropsci2005.0797.
Bradbury, P.J., Z. Zhang, D.E. Kroon, T.M. Casstevens, Y. Ramdoss and E.S. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633-2635. doi:10.1093/bioinformatics/btm308.
Brazauskas, G., I. Lenk, M.G. Pedersen, B. Studer and T. Lübberstedt. 2011. Genetic variation, population structure, and linkage disequilibrium in European elite germplasm of perennial ryegrass. Plant Science 181: 412-420. doi:10.1016/j.plantsci.2011.06.013.
Brazauskas, G., I. Pašakinskienė, T. Asp and T. Lübberstedt. 2010. Nucleotide diversity and linkage disequilibrium in five *Lolium perenne* genes with putative role in shoot morphology. Plant Science 179: 194-201. doi:10.1016/j.plantsci.2010.04.016.
Breseghello, F. and A.S.G. Coelho. 2013. Traditional and modern plant breeding methods with examples in rice (*Oryza sativa* L.). Journal of Agricultural and Food Chemistry 61: 8277-8286. doi:10.1021/jf305531j.
Bretting, P.K. 2018. 2017 Frank Meyer Medal for Plant Genetic Resources Lecture: Stewards of Our Agricultural Future. Crop Science 58: 2233-2240. doi:10.2135/cropsci2018.05.0334.
Brewbaker, J.L. 2008. Registration of KX2-Hawaii, interspecific-hybrid *Leucaena*. Journal of Plant Registrations 2: 190-193. doi:10.3198/jpr2007.05.0298crc.
Brewbaker, J.L. and W.F. Keim. 1953. A fertile interspecific hybrid in *Trifolium* (4n *T. repens* L. x 4n *T. nigrescens* Viv.). The American Naturalist 87: 323-326. doi:10.1086/281790.
Briggs, W.H. and I.L. Goldman. 2006. Genetic variation and selection response in model breeding populations of *Brassica rapa* following a diversity bottleneck. Genetics 172: 457-465. doi:10.1534/genetics.105.040899.
Brock, J. 1988. Evaluation of New Zealand bred white clover cultivars under rotational grazing and set stocking with sheep. Proceedings of the New Zealand Grassland Association 49: 203-206.
Brock, J. and M. Hay. 1996. A review of the role of grazing management on the growth and performance of white clover cultivars in lowland New Zealand pastures. Special Publication - Agronomy Society of New Zealand 11: 65-70.
Brock, J. and M. Hay. 2001. White clover performance in sown pastures: A biological/ecological perspective. Proceedings of the New Zealand Grassland Association 63: 73-84.
Brock, J., M. Hyslop and K. Widdup. 2003. A review of red and white clovers in the dryland environment. Legumes for dryland pastures. DJ Moot (Ed) Grassland Research and Practices 11: 101-107.
Brock, J. and J. Tilbrook. 2000. Effect of cultivar of white clover on plant morphology during the establishment of mixed pastures under sheep grazing. New Zealand Journal of Agricultural Research 43: 335-343.
Brock, J.R., J.L. Caradus and M.J.M. Hay. 1989. Fifty Years of White Clover Research in New Zealand. Proceedings of the New Zealand Grassland Association 50: 25-39.
Brown, A. 1989a. Core collections: a practical approach to genetic resources management. Genome 31: 818-824. doi:10.1139/g89-144.

Brown, A. 1989b. The case for core collections. In: A. Brown, O. Frankel, D. Marshall and J. Williams, editors, The Use of Plant Genetic Resources. Cambridge University Press, Cambridge. p. 136-156.
Brown, C. and R. Green. 2003. The challenges facing legumes in a dryland environment-a consultant's view. Legumes for dryland pastures. Grassland Research and Practice Series 11: 7-12.
Brown, H., D. Moot and K. Pollock. 2005. Herbage production, persistence, nutritive characteristics and water use of perennial forages grown over 6 years on a Wakanui silt loam. New Zealand Journal of Agricultural Research 48: 423-439.
Bruce, R.W., D. Torkamaneh, C. Grainger, F. Belzile, M. Eskandari and I. Rajcan. 2019. Genome-wide genetic diversity is maintained through decades of soybean breeding in Canada. Theoretical and Applied Genetics 132: 3089-3100. doi:10.1007/s00122-019-03408-y.
Brummer, E. 2013. Global impact of sown temperate pastures on productivity and ecosystem stability - what progress have we made? Revitalising Grasslands to Sustain our Communities: Proceedings, 22nd International Grassland Congress: 276-281.
Brummer, E.C. 1999. Capturing heterosis in forage crop cultivar development. Crop Science 39: 943-954.
Bryant, W.G. 1974. Caucasian clover (*Trifolium ambiguum* Bieb.): A review. The Journal of the Australian Institute of Agricultural Science 40: 11-19.
Bulińska-Radomska, Z. 2000. Morphological relationships among 15 species of *Trifolium* occurring in Poland. Genetic Resources and Crop Evolution 47: 267-272. doi:10.1023/a:1008707024828.
Burdon, J. 1980. Intra-specific diversity in a natural population of *Trifolium repens*. The Journal of Ecology: 717-735. doi:10.2307/2259452.
Busbice, T.H. and C. Wilsie. 1966. Inbreeding depression and heterosis in autotetraploids with application to *Medicago sativa* L. Euphytica 15: 52-67.
Byrne, P.F., G.M. Volk, C. Gardner, M.A. Gore, P.W. Simon and S. Smith. 2018. Sustaining the future of plant breeding: The critical role of the USDA-ARS National Plant Germplasm System. Crop Science 58: 451-468. doi:10.2135/cropsci2017.05.0303.
Byrne, S., A. Czaban, B. Studer, F. Panitz, C. Bendixen and T. Asp. 2013. Genome-wide allele frequency fingerprints (GWAFFs) of populations via genotyping by sequencing. PloS One 8: e57438. doi:10.1371/journal.pone.0057438.
Calboli, F.C., J. Sampson, N. Fretwell and D.J. Balding. 2008. Population structure and inbreeding from pedigree analysis of purebred dogs. Genetics 179: 593-601. doi:10.1534/genetics.107.084954.
1994. Variation in the spread of white clover plants growing in competition with different grasses. Proceedings Agronomy Society of NZ.
Caradus, J. 1986. World checklist of white clover varieties. New Zealand Journal of Experimental Agriculture 14: 119-164. doi:10.1080/03015521.1986.10426137.
Caradus, J. 1994. Genetic diversity within white clover (*Trifolium repens* L.). Proceedings of the Agronomy Society of New Zealand 24: 2.
Caradus, J., J. Brock and M. Hay. 1989a. Fifty years of white clover research in New Zealand. Proceedings of the New Zealand Grassland Association 50: 25-39.
Caradus, J. and D. Chapman. 1996a. Selection for and heritability of stolon characteristics in two cultivars of white clover. Crop Science 36: 900-904. doi:10.2135/cropsci1996.0011183X0036000400014x.
Caradus, J. and B. Christie. 1998a. Winter hardiness and artificial frost tolerance of white clover ecotypes and selected breeding lines. Canadian Journal of Plant Science 78: 251-255. doi:10.4141/P96-147.
Caradus, J., P. Clifford, D. Chapman, G. Cousins, W. Williams and J. Miller. 1997a. Breeding and description of 'Grasslands Sustain', a medium-large-leaved white clover (*Trifolium repens* L.) cultivar. New Zealand Journal of Agricultural Research 40: 1-7.
Caradus, J., M. Hay, A. Mackay, V. Thomas, J. Dunlop, M. Lambert, et al. 1993a. Variation within white clover (*Trifolium repens* L.) for phenotypic plasticity of morphological and yield related characters, induced by phosphorus supply. New Phytologist 123: 175-184. doi:10.1111/j.1469-8137.1993.tb04543.x.
Caradus, J., R. Hay and D. Woodfield. 1996b. The positioning of white clover cultivars in New Zealand. Special Publication - Agronomy Society of New Zealand: 45-50.

Caradus, J. and A. Mackay. 1989b. Morphological and flowering variation of *Trifolium dubium* Sibth. New Zealand Journal of Agricultural Research 32: 129-132.
Caradus, J. and W. Williams. 1989c. Breeding for legume persistence in New Zealand. In: G. C. Marten, A. G. Matches, R. F. Barnes, R. W. Brougham, R. J. Clements and G. W. Sheath, editors, Persistence of Forage Legumes. ASA/CSSA/SSSA, Madison, Wisconsin, USA. p. 523-539.
Caradus, J. and D. Woodfield. 1997b. World checklist of white clover varieties II. New Zealand Journal of Agricultural Research 40: 115-206. doi:10.1080/00288233.1997.9513239.
Caradus, J. and D. Woodfield. 1998b. Genetic control of adaptive root characteristics in white clover. Plant and Soil 200: 63-69.
Caradus, J., D. Woodfield and A. Stewart. 1996c. Overview and vision for white clover. Special Publication - Agronomy Society of New Zealand: 1-6.
Caradus, J.R. 1981. Root growth of white clover (*Trifolium repens* L.) lines in glass–fronted containers. New Zealand Journal of Agricultural Research 24: 43-54. doi:10.1080/00288233.1981.10420870.
Carlsson, G. and K. Huss-Danell. 2003. Nitrogen fixation in perennial forage legumes in the field. Plant and Soil 253: 353-372. doi:10.1023/A:1024847017371.
Casler, M. 1995. Patterns of variation in a collection of perennial ryegrass accessions. Crop Science 35: 1169-1177. doi:10.2135/cropsci1995.0011183X003500040043x.
Casler, M., J.F. Pedersen, G. Eizenga and S. Stratton. 1996. Germplasm and cultivar development. Agronomy & Horticulture - Faculty Publications 954.
Casler, M.D. 1998. Genetic variation within eight populations of perennial forage grasses. Plant Breeding 117: 243-249. doi:10.1111/j.1439-0523.1998.tb01933.x.
Casler, M.D. and E.C. Brummer. 2008. Theoretical expected genetic gains for among-and-within-family selection methods in perennial forage crops. Crop Science 48: 890-902. doi:10.2135/cropsci2007.09.0499.
Cassileth, B. 2010. Red clover (*Trifolium pratense*). Oncology 24: 960.
Catchen, J., P.A. Hohenlohe, S. Bassham, A. Amores and W.A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Molecular Ecology 22: 3124-3140. doi:10.1111/mec.12354.
Cattivelli, L., F. Rizza, F.-W. Badeck, E. Mazzucotelli, A.M. Mastrangelo, E. Francia, et al. 2008. Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. Field Crops Research 105: 1-14. doi: 10.1016/j.fcr.2007.07.004.
Cervantes, I., A. Molina, F. Goyache, J.P. Gutiérrez and M. Valera. 2008. Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. Livestock Science 113: 24-33. doi:10.1016/j.livsci.2007.02.011.
Chapman, D., J. Lee, L. Rossi, G. Edwards, J. Pinxterhuis and E. Minnee. 2017. White clover: the forgotten component of high-producing pastures? Animal Production Science 57: 1269-1276. doi:10.1071/AN16453.
Chapman, D.F. 1983. Growth and demography of *Trifolium repens* stolons in grazed hill pastures. Journal of Applied Ecology 20: 597-608. doi:10.2307/2403529.
Chappell, P.R. 2015. The climate and weather of Manawatu-Wanganui.
Charlson, D.V., S. Bhatnagar, C.A. King, J.D. Ray, C.H. Sneller, T.E. Carter, et al. 2009. Polygenic inheritance of canopy wilting in soybean [*Glycine max* (L.) Merr.]. Theoretical and Applied Genetics 119: 587-594. doi:10.1007/s00122-009-1068-4.
Charlton, D. and A. Stewart. 2006. Pasture and forage plants for New Zealand. New Zealand Grassland Trust.
Charlton, J. and N. Giddens. 1983. Establishment of hill country white clover selections from oversowing. Proceedings of the New Zealand Grassland Association 44: 149-155. doi:10.1080/00288233.1981.10420870.
Charlton, J. and A. Stewart. 1999. Pasture species and cultivars used in New Zealand - a list. Proceedings of the Conference - New Zealand Grassland Association: 147-166.
Charlton, J., D. Woodfield and J. Caradus. 1989. Performance of *Trifolium repens* genotypes under grazing in New Zealand hill pasture. Proceedings of the XVI International Grasslands Congress: 349-350.
Chen, J.-Q., X.-P. Meng, Y. Zhang, M. Xia and X.-P. Wang. 2008. Over-expression of OsDREB genes lead to enhanced drought tolerance in rice. Biotechnology letters 30: 2191-2198. doi:10.1007/s10529-008-9811-5.

Chen, J., R. Chopra, C. Hayes, G. Morris, S. Marla, J. Burke, et al. 2017. Genome-wide association study of developing leaves' heat tolerance during vegetative growth stages in a sorghum association panel. The Plant Genome 10. doi:10.3835/plantgenome2016.09.0091.
Choo, T. 1988. Plant regeneration in zigzag clover (*Trifolium medium* L.). Plant Cell Reports 7: 246-248. doi:10.1007/BF00272534.
Chou, M.-c. and P.B. Gibson. 1968. Cross-compatibility of *Trifolium nigrescent* with diploid and tetraploid *Trifolium occidentale*. Crop Science 8: 266-267.
Clark, R.L., H.L. Shands, P.K. Bretting and S.A. Eberhart. 1997. Managing large diverse germplasm collections. Crop Science 37: 1-6. doi:10.2135/cropsci1997.0011183X003700010001x.
Claydon, R.B., W. Rumball and J.E. Miller. 2003. 'Grasslands Sensation' red clover (*Trifolium pratense* L.). New Zealand Journal of Agricultural Research 46: 355-357. doi:10.1080/00288233.2003.9513564.
Cleveland, R.W. 1985. Reproductive cycle and cytogenetics. Clover Science and Technology: 71-110.
Cnops, G., A. Rohde, O. Saracutu, M. Malengier and I. Roldán-Ruiz. 2010. Morphological and molecular diversity of branching in red clover (*Trifolium pratense*). Sustainable Use of Genetic Diversity in Forage and Turf Breeding. Springer. p. 73-77.
Cogan, N., M.T. Abberton, K. Smith, G. Kearney, A.H. Marshall, A. Williams, et al. 2006. Individual and multi-environment combined analyses identify QTLs for morphogenetic and reproductive development traits in white clover (*Trifolium repens* L.). Theoretical and Applied Genetics 112: 1401-1415. doi:10.1007/s00122-006-0241-2.
Cole, J.B. 2007. PyPedal: A computer program for pedigree analysis. Computers and Electronics in Agriculture 57: 107-113.
Collins, R.P. 1998. The effect of drought stress and winter stress on the persistence of white clover. Proceedings of FAO/CIHEAM Lowland Grasslands Sub-Network Meeting.
Collins, R.P., Á. Helgadóttir, M. Fothergill and I. Rhodes. 2002. Variation amongst survivor populations of white clover collected from sites across Europe: Growth attributes and physiological responses to low temperature. Annals of Botany 89: 283-292. doi:10.1093/aob/mcf037.
Collins, R.P., Á. Helgadóttir, B.E. Frankow-Lindberg, L. Skøt, C. Jones and K.P. Skøt. 2012. Temporal changes in population genetic diversity and structure in red and white clover grown in three contrasting environments in northern Europe. Annals of Botany 110: 1341-1350. doi:10.1093/aob/mcs058.
Comai, L. 2005. The advantages and disadvantages of being polyploid. Nature Reviews Genetics 6: 836. doi:10.1038/nrg1711.
Conaty, W.C., J.R. Mahan, J.E. Neilsen and G.A. Constable. 2014. Vapour pressure deficit aids the interpretation of cotton canopy temperature response to water deficit. Functional Plant Biology 41: 535-546. doi:10.1071/FP13223.
Connolly, V. 1990. Seed yield and yield components in ten white clover cultivars. Irish Journal of Agricultural Research: 41-48.
Cooper, B. and D. Chapman. 1993. Grasslands Prestige (G39), a white clover cultivar originating from northern New Zealand. Proceedings of the International Grassland Congress 17: 458-459.
Cooper, B., P. Clifford and W. Williams. 1997. Development of white clover (*Trifolium repens* L.) cultivar Grasslands Challenge (G23). Proceedings of the Conference - New Zealand Grassland Association: 99-102.
Corkill, L. 1945. Short rotation ryegrass, its breeding and characteristics. New Zealand Journal of Agricultural Research 71: 465-470.
Corkill, L. 1949. Pasture improvement in New Zealand. Experimental Agriculture 17: 157-169.
Coster, A. 2015. Package 'pedigree'. doi:cran.r-project.org/web/packages/pedigree/pedigree.pdf.
Crossa, J., G. de Los Campos, P. Pérez, D. Gianola, J. Burgueño, J.L. Araus, et al. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186: 713-724. doi:10.1534/genetics.110.118521.
Crow, J.F. and M. Kimura. 1970a. An introduction to population genetics theory. Harper & Row, Publishers, New York, Evanston and London.
Crush, J., L. Ouyang and S. Nichols. 2015. Root morphology and architecture, and internal phosphate use efficiency, in related white clover cultivars of different ages. New Zealand Journal of Agricultural Research 58: 302-310. doi:10.1080/00288233.2015.1029075.

Crush, J.R. 1987. Nitrogen fixation. In: M. J. Baker and W. M. Williams, editors, White Clover. CAB International, Wallingford, U.K. p. 185-202.
D'Agostino, N., F. Taranto, S. Camposeo, G. Mangini, V. Fanelli, S. Gadaleta, et al. 2018. GBS-derived SNP catalogue unveiled wide genetic variability and geographical relationships of Italian olive cultivars. Scientific Reports 8: 1-13. doi:10.1038/s41598-018-34207-y.
Dabney, A.S., J.D; Warnes G.R. 2010. qvalue: Q-value estimation for false discovery rate control. R package, version 1.0.
Daday, H. 1954a. Gene frequencies in wild populations of *Trifolium repens* L. I Distribution by latitude. Heredity 8: 61-78.
Daday, H. 1954b. Gene frequencies in wild populations of *Trifolium repens* L. II. Distribution by altitude. Heredity 8: 377-384.
Daday, H. 1965. Gene frequencies in wild populations of *Trifolium repens* L. IV. Mechanism of natural selection. Heredity 20: 355-365.
Daetwyler, H.D., A.A. Swan, J.H. van der Werf and B.J. Hayes. 2012. Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. Genetics Selection Evolution 44: 33. doi:10.1186/1297-9686-44-33.
Daly, G. and C. Mason. 1987. Performance of Caucasian and zigzag clovers. Proceedings of the New Zealand Grassland Association 48: 151-156.
Danecek, P., A. Auton, G. Abecasis, C.A. Albers, E. Banks, M.A. DePristo, et al. 2011. The variant call format and VCFtools. Bioinformatics 27: 2156-2158. doi:10.1093/bioinformatics/btr330.
Davies, A. and D.R. Jones. 1992. The production of leaves and stolon branches on established white clover cuttings in relation to temperature and soil moisture in the field. Annals of Botany 69: 515-521.
de Colmenares, N.G., J.R. Ramírez-Martínez, J.O. Aldana, M.E. Ramos-Niño, M.N. Clifford, S. Pékerar, et al. 1998. Isolation, characterisation and determination of biological activity of coffee proanthocyanidins. Journal of the Science of Food and Agriculture 77: 368-372. doi:10.1002/(SICI)1097-0010(199807)77:3<368::AID-JSFA52>3.0.CO;2-V.
De Vega, J.J., S. Ayling, M. Hegarty, D. Kudrna, J.L. Goicoechea, Å. Ergon, et al. 2015. Red clover (*Trifolium pratense* L.) draft genome provides a platform for trait improvement. Scientific reports 5: 17394. doi:10.1038/srep17394.
Dear, B. and M. Zorin. 1985. Persistence and productivity of *Trifolium ambiguum* M. Bieb.(Caucasian clover) in a high altitude region of south-eastern Australia. Australian Journal of Experimental Agriculture 25: 124-132. doi:10.1071/EA9850124.
Dessureaux, L. and A. Gallais. 1969. Inbreeding and heterosis in autotetraploid alfalfa. I. Fertility. Canadian Journal of Genetics and Cytology 11: 706-715.
Devi, M.J., T.R. Sinclair and E. Taliercio. 2015. Comparisons of the effects of elevated vapor pressure deficit on gene expression in leaves among two fast-wilting and a slow-wilting soybean. PloS One 10. doi:10.1371/journal.pone.0139134.
Dias, P.M.B., B. Julier, J.-P. Sampoux, P. Barre and M. Dall'Agnol. 2008. Genetic diversity in red clover (*Trifolium pratense* L.) revealed by morphological and microsatellite (SSR) markers. Euphytica 160: 189-205. doi:10.1007/s10681-007-9534-z.
Díez, M.J., L. De la Rosa, I. Martín, L.M. Guasch Pereira, M.E. Cartea, C. Mallor, et al. 2018. Plant genebanks: present situation and proposals for their improvement. The case of the Spanish network. Frontiers in Plant Science 9: 1794. doi:10.3389/fpls.2018.01794.
Dirzo, R. and J.L. Harper. 1982. Experimental studies on slug-plant interactions: III. Differences in the acceptability of individual plants of *Trifolium repens* to slugs and snails. The Journal of Ecology: 101-117.
Dixon, R.A., D.Y. Xie and S.B. Sharma. 2005. Proanthocyanidins–a final frontier in flavonoid research? New phytologist 165: 9-28. doi:10.1111/j.1469-8137.2004.01217.x.
Dluhošová, J., J. Ištvánek, J. Nedělník and J. Řepková. 2018. Red clover (*Trifolium pratense*) and zigzag clover (*T. medium*)–a picture of genomic similarities and differences. Frontiers in Plant Science 9: 724. doi:10.3389/fpls.2018.00724.
Dodd, M., G. Sheath and S. Richardson. 1995a. Development of subterranean clover (*Trifolium subterraneum* L.) genotypes for New Zealand pastures: 1. Whatawhata persistence evaluation. New Zealand Journal of Agricultural Research 38: 33-47.

Dodd, M., G. Sheath and S. Richardson. 1995b. Development of subterranean clover (*Trifolium subterraneum* L.) genotypes for New Zealand pastures: 2. Wairakei persistence evaluation. New Zealand Journal of Agricultural Research 38: 49-56.
Dodd, M., G. Sheath and I. Tarbotton. 1995c. Development of subterranean clover (*Trifolium subterraneum* L.) genotypes for New Zealand pastures: 3. Whatawhata production evaluation. New Zealand Journal of Agricultural Research 38: 57-63.
Dodds, K.G., J.C. McEwan, R. Brauning, R.M. Anderson, T.C. van Stijn, T. Kristjánsson, et al. 2015. Construction of relatedness matrices using genotyping-by-sequencing data. BMC Genomics 16: 1047. doi:10.1186/s12864-015-2252-3.
Dolstra, O., C. Denneboom, L.F.d. Vos and E.N.V. Loo. 2007. Marker-assisted selection for improving quantitative traits of forage crops. In: E. P. Guimarães, editor Marker-assisted Selection: Current Status and Future Perspectives in Crops, Livestock, Forestry and Fish. Food and Agriculture Organization of the United Nations.
Drach, N., V. Rogovchenko, T. Skorobagat'ko and G. Efimenko. 1986. Significance of self pollination in the reproduction of early red clover. Sb. Nauchnykh Tr. Prikl. Bot. Genet. Sel 103: 31-34.
Dreisigacker, S., P. Zhang, M. Warburton, M. Van Ginkel, D. Hoisington, M. Bohn, et al. 2004. SSR and pedigree analyses of genetic diversity among CIMMYT wheat lines targeted to different megaenvironments. Crop Science 44: 381-388. doi:10.2135/cropsci2004.0381.
Du, Q., W. Lu, M. Quan, L. Xiao, F. Song, P. Li, et al. 2018. Genome-wide association studies to improve wood properties: challenges and prospects. Frontiers in Plant Science 9: 1912. doi:10.3389/fpls.2018.01912.
Dudchenko, O., M. Pham, C. Lui, S.S. Batra, M. Hoeger, S.K. Nyquist, et al. 2018. Hi-C yields chromosome-length scaffolds for a legume genome, *Trifolium subterraneum*. bioRxiv. doi:10.1101/473553.
Duncan, D.B. 1955. Multiple range and multiple F tests. Biometrics 11: 1-42.
Duncan, J., M. Anderson and N. Taylor. 1973. Effect of inbreeding on pseudo-self-compatibility in red clover (*Trifolium pratense* L.). Euphytica 22: 535-542. doi:10.1007/BF00036653.
Durand, J.-L., F. Gastal, S. Etchebest, A.-C. Bonnet and M. Ghesquière. 1997. Interspecific variability of plant water status and leaf morphogenesis in temperate forage grasses under summer water deficit. Developments in Crop Science. Elsevier. p. 135-143.
Dwivedi, S.L., A. Scheben, D. Edwards, C. Spillane and R. Ortiz. 2017. Assessing and exploiting functional diversity in germplasm pools to enhance abiotic stress adaptation and yield in cereals and food legumes. Frontiers in Plant Science 8: 1461. doi:10.3389/fpls.2017.01461.
Egan, L.M., R.W. Hofmann, B.A. Barrett, K. Ghamkhar and V. Hoyos-Villegas. 2019a. Identification of founding accessions and patterns of relatedness and inbreeding derived from historical pedigree data in a white clover germplasm collection in New Zealand. Crop Science 59: 2087-2099. doi:10.2135/cropsci2018.11.0688.
Egan, L.M., R.W. Hofmann, K. Ghamkhar and V. Hoyos-Villegas. 2019b. Identification of founding accessions and patterns of relatedness and inbreeding derived from historical pedigree data in a red clover germplasm collection in New Zealand. Crop Science 59: 2100-2108. doi:10.2135/cropsci2019.01.0045.
Egan, L.M., R.W. Hofmann, P. Seguin, K. Ghamkhar and V. Hoyos-Villegas. 2020. Pedigree analysis of pre-breeding efforts in *Trifolium* spp. germplasm in New Zealand. BMC Genetics.
El-Kassaby, Y.A., E.P. Cappa, C. Liewlaksaneeyanawin, J. Klápště and M. Lstibůrek. 2011. Breeding without breeding: is a complete pedigree necessary for efficient breeding? PloS One 6: e25737. doi:10.1371/journal.pone.0025737.
Ellis, W. and N.R. Young. 1967. The characteristics of European, Mediterranean and other populations of white clover (*Trifolium repens* L.). Euphytica 16: 330-340. doi:10.1007/BF00028939.
Ellison, N.W., A. Liston, J.J. Steiner, W.M. Williams and N.L. Taylor. 2006. Molecular phylogenetics of the clover genus (*Trifolium*—Leguminosae). Molecular Phylogenetics and Evolution 39: 688-705. doi:10.1016/j.ympev.2006.01.004.
Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, et al. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One 6: e19379. doi:10.1371/journal.pone.0019379.
Falconer, D.S. 1975. Introduction to quantitative genetics. Pearson Education India.
FAO. 2010. Commission on genetic resources for food and agriculture.

Faville, M., A. Griffiths, M. Jahufer and B. Barrett. 2012. Progress towards marker-assisted selection in forages. Proceedings of the New Zealand Grassland Association 74: 189-194.
Fè, D., F. Cericola, S. Byrne, I. Lenk, B.H. Ashraf, M.G. Pedersen, et al. 2015. Genomic dissection and prediction of heading date in perennial ryegrass. BMC Genomics 16: 921. doi:10.1186/s12864-015-2163-3.
Fehr, W.R. 1991. Principles of Cultivar Development: Theory and Technique. 1 ed. Macmillan Publishing Company.
Ferguson, N., E. Rupert and P. Evans. 1990. Interspecific *Trifolium* hybrids produced by embryo and ovule culture. Crop Science 30: 1145-1149. doi:10.2135/cropsci1990.0011183X003000050039x.
Fernando, R. and D. Habier. 2006. Kinship and Inbreeding. eLS. doi:10.1038/npg.els.0005398.
Fischlin, A., G.F. Midgley, J.T. Price, R. Leemans, B. Gopal, C.M. Turley, et al. 2007. Ecosystems, their properties goods and services. In: M. L. Parry, J. P. Canziani, J. P. Palutikof, P. J. van der Linden and C. E. Hanson, editors, Climate change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge University Press, Cambridge, UK. p. 211-272.
Fletcher, A.L., T.R. Sinclair and L.H. Allen Jr. 2007. Transpiration responses to vapor pressure deficit in well watered 'slow-wilting'and commercial soybean. Environmental and Experimental Botany 61: 145-151. doi:10.1016/j.envexpbot.2007.05.004.
Fleury, D., S. Jefferies, H. Kuchel and P. Langridge. 2010. Genetic and genomic tools to improve drought tolerance in wheat. Journal of Exerimental Botany 61: 3211-3222. doi:10.1093/jxb/erq152.
Ford, J. and B. Barrett. 2011. Improving red clover persistence under grazing. Proceedings of the New Zealand Grassland Association 73: 119-124.
Ford, J., G. Cousins, M. Jahufer, I. Baird, D. Woodfield and B. Barrett. 2015. Grasslands Legacy–a new, large-leafed white clover cultivar with broad adaption. Journal of New Zealand Grasslands 77: 211-217.
Fradgley, N., K.A. Gardner, J. Cockram, J. Elderfield, J.M. Hickey, P. Howell, et al. 2019. A large-scale pedigree resource of wheat reveals evidence for adaptation and selection by breeders. PloS Biology 17: e3000071. doi:10.1371/journal.pbio.3000071.
Frame, J. 2019. Forage legumes for temperate grasslandsCRC Press.
Frame, J. and P. Newbould. 1986. Agronomy of White Clover. In: N. C. Brady, editor Advances in Agronomy. Academic Press. p. 1-88.
Francis, C., J. Gladstones and W. Stern. 1970. Selection of new subterranean clover cultivars in southwestern Australia. Proceedings of the 11th Grasslands Congress: 214-218.
Frankel, O.H. and A. Brown. 1984. Plant genetic resources today: a critical appraisal. Crop Genetic Resources: Conservation and Evaluation.
Fu, S., M. Hill and J. Hampton. 2001. Root system development in Caucasian clover cv. Monaro and its contribution to seed yield. New Zealand Journal of Agricultural Research 44: 23-29. doi:10.1080/00288233.2001.9513458.
Fu, Y.-B. 2015. Understanding crop genetic diversity under modern plant breeding. Theoretical and Applied Genetics 128: 2131-2142. doi:10.1007/s00122-015-2585-y.
Gage, J.L., D. Jarquin, C. Romay, A. Lorenz, E.S. Buckler, S. Kaeppler, et al. 2017. The effect of artificial selection on phenotypic plasticity in maize. Nature Communications 8: 1-11. doi:10.1038/s41467-017-01450-2.
Gasparis, S., W. Orczyk, W. Zalewski and A. Nadolska-Orczyk. 2011. The RNA-mediated silencing of one of the Pin genes in allohexaploid wheat simultaneously decreases the expression of the other, and increases grain hardness. Journal of Experimental Botany 62: 4025-4036. doi:10.1093/jxb/err103.
Gautier, H., C. Varlet-Grancher and J.M. Membré. 2001. Plasticity of petioles of white clover (*Trifolium repens*) to blue light. Physiologia Plantarum 112: 293-300. doi:10.1034/j.1399-3054.2001.1120219.x.
George, J., M.P. Dobrowolski, E. van Zijll de Jong, N.O. Cogan, K.F. Smith and J.W. Forster. 2006. Assessment of genetic diversity in cultivars of white clover (*Trifolium repens* L.) detected by SSR polymorphisms. Genome 49: 919-930. doi:10.1139/g06-079.

George, R.M. 2014. Testing alternative breeding methods in white clover. Ph.D Lincoln University, Christchurch, New Zealand.
Gepts, P. 2004. Crop domestication as a long-term selection experiment. Plant Breeding Reviews 24: 1-44. doi:10.1002/9780470650288.ch1.
Gerard, P., C. Ferguson and S. van Amsterdam. 2017. Comparison of New Zealand perennial clovers for resilience against common pasture pests. New Zealand Plant Protection 70: 241-249. doi:10.30843/nzpp.2017.70.57.
Ghamkhar, K., S. Isobe, P.G. Nichols, T. Faithfull, M.H. Ryan, R. Snowball, et al. 2012. The first genetic maps for subterranean clover (*Trifolium subterraneum* L.) and comparative genomics with *T. pratense* L. and *Medicago truncatula* Gaertn. to identify new molecular markers for breeding. Molecular Breeding 30: 213-226. doi:10.1007/s11032-011-9612-8.
Ghimiray, M. and R. Vernooy. 2017. The importance and challenges of crop germplasm interdependence: the case of Bhutan. Food Security 9: 301-310. doi:10.1007/s12571-017-0647-5.
Gholipoor, M., T. Sinclair, M. Raza, C. Löffler, M. Cooper and C. Messina. 2013. Maize hybrid variability for transpiration decrease with progressive soil drying. Journal of Agronomy and Crop Science 199: 23-29. doi:10.1111/j.1439-037X.2012.00530.x.
Gibson, G. 2012. Rare and common variants: twenty arguments. Nature Reviews Genetics 13: 135-145. doi:10.1038/nrg3118.
Gibson, P.B. and G. Beinhart. 1969. Hybridization of *Trifolium occidentale* With two other species of clover. Journal of Heredity 60: 93-96.
Gillard, P., A. Bishop and R. Reid. 1989. Introduction and evaluation of pasture legumes in high rainfall north-western Tasmania. Proceedings of the 5th Australian Agronomy Conference 521.
Gillett, J.M., N.L. Taylor and M. Collins. 2001. World of Clovers, Iowa State University Press.
Gilmour, A.R., B.J. Gogel, B.R. Cullis, S.J. Welham and R. Thompson. 2015. ASReml User Guide Release 4.1 Functional Specification. VSN International Ltd Hemel Hempstead, HP1 1ES, UK.
Gizlice, Z., T. Carter and J. Burton. 1994. Genetic base for North American public soybean cultivars released between 1947 and 1988. Crop Science 34: 1143-1151. doi:10.2135/cropsci1994.0011183X003400050001x.
González, J.d.J.S., J.A.R. Corral, G.M. García, G.R. Ojeda, L. De la Cruz Larios, J.B. Holland, et al. 2018. Ecogeography of teosinte. PloS One 13: e0192676. doi:10.1371/journal.pone.0192676.
Gorjanc, G., J. Jenko, S.J. Hearne and J.M. Hickey. 2016. Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. BMC Genomics 17: 1. doi:10.1186/s12864-015-2345-z.
Graczyk, M., K. Andres, E. Kapkowska and T. Szwaczkowski. 2015. Pedigree analyses of the Zatorska goose population. Czech Journal of Animal Science 60: 513-520. doi:10.17221/8560-CJAS.
Graner, A., W.F. Ludwig and A.E. Melchinger. 1994. Relationships among European barley germplasm: II. Comparison of RFLP and pedigree data. Crop Science 34: 1199-1205.
Grasslands Division, D.S.I.R., New Zealand. 1971. Register of Australian Herbage Plant Cultivars: B. Legumes. 1. Clover. *Trifolium pratense* L. (red clover) cv. Grasslands Hamua.
GrasslanzTechnology. Grasslanz Product Guide.
Graudal, L., E.D. Kjær and S. Canger. 1995. A systematic approach to the conservation of genetic resources of trees and shrubs in Denmark. Forest Ecology and Management 73: 117-134.
Griffiths, A.G., B.A. Barrett, D. Simon, A.K. Khan, P. Bickerstaff, C.B. Anderson, et al. 2013. An integrated genetic linkage map for white clover (*Trifolium repens* L.) with alignment to Medicago. BMC Genomics 14: 388.
Griffiths, A.G., R. Moraga, M. Tausen, V. Gupta, T.P. Bilton, M.A. Campbell, et al. 2019. Breaking free: the genomics of allopolyploidy-facilitated niche expansion in white clover. The Plant Cell 31: 1466-1487. doi:10.1105/tpc.18.00606.
Gupta, P.K., P.L. Kulwal and V. Jaiswal. 2019. Association mapping in plants in the post-GWAS genomics era. Advances in genetics. Elsevier. p. 75-154.
Gustine, D.L. and M.A. Sanderson. 2001. Molecular analysis of white clover population structure in grazed swards during two growing seasons. Crop Science 41: 1143-1149. doi:10.2135/cropsci2001.4141143x.

Gustine, D.L., P.W. Voigt, E.C. Brummer and Y.A. Papadopoulos. 2002. Genetic variation of RAPD markers for North American white clover collections and cultivars. Crop Science 42: 343-347. doi:10.2135/cropsci2002.3430.
Gutierrez-Gil, B., J.J. Arranz, R. Pong-Wong, E. García-Gámez, J. Kijas and P. Wiener. 2014. Application of selection mapping to identify genomic regions associated with dairy production in sheep. PloS One 9. doi:10.1371/journal.pone.0094623.
Hagenblad, J. and M. Nordborg. 2002. Sequence variation and haplotype structure surrounding the flowering time locus FRI in *Arabidopsis thaliana*. Genetics 161: 289-298.
Halewood, M., T. Chiurugwi, R. Sackville Hamilton, B. Kurtz, E. Marden, E. Welch, et al. 2018. Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution. New Phytologist 217: 1407-1419. doi:10.1111/nph.14993.
Hallauer, A.R. 1992. Recurrent selection in maize. Plant Breeding Reviews 9: 115-179. doi:10.1002/9780470650363.ch6.
Hamann, H. and O. Distl. 2008. Genetic variability in Hanoverian warmblood horses using pedigree analysis. Journal of Animal Science 86: 1503-1513. doi:10.2527/jas.2007-0382.
Han, K., H.Y. Lee, N.Y. Ro, O.S. Hur, J.H. Lee, J.K. Kwon, et al. 2018. QTL mapping and GWAS reveal candidate genes controlling capsaicinoid content in Capsicum. Plant Biotechnology. doi:10.1111/pbi.12894.
Hancock, K.R., V. Collette, K. Fraser, M. Greig, H. Xue, K. Richardson, et al. 2012. Expression of the R2R3-MYB transcription factor TaMYB14 from *Trifolium arvense* activates proanthocyanidin biosynthesis in the legumes *Trifolium repens* and *Medicago sativa*. Plant Physiology 159: 1204-1220. doi:10.1104/pp.112.195420.
Hand, M.L., N.O. Cogan and J.W. Forster. 2012. Genome-wide SNP identification in multiple morphotypes of allohexaploid tall fescue (*Festuca arundinacea* Schreb). BMC Genomics 13: 219. doi:10.1186/1471-2164-13-219.
Hanif, M., F. Gao, J. Liu, W. Wen, Y. Zhang, A. Rasheed, et al. 2016. TaTGW6-A1, an ortholog of rice TGW6, is associated with grain weight and yield in bread wheat. Molecular Breeding 36: 1. doi:10.1007/s11032-015-0425-z.
Harmer, M., A. Stewart and D. Woodfield. 2016. Genetic gain in perennial ryegrass forage yield in Australia and New Zealand. Journal of New Zealand Grasslands 78: 133-138.
Harris, A., A. Chu, R. Burgess and J. Brock. 1985. Limitations to production and choice of species in finishing pastures. Using herbage cultivars. Grassland Research and Practice Series (3): 53-57.
Harris, S., D. Clark, M. Auldist, C. Waugh and P. Laboyrie. 1997. Optimum white clover content for dairy pastures. New Zealand Grassland Association: 29-34.
Hawkes, J.G. 1977. The importance of wild germplasm in plant breeding. Euphytica 26: 615-621. doi:10.1007/bf00021686.
Hay, R. and D. Ryan. 1989. A review of 10 years' research with red clovers under grazing in Southland. Proceedings of the New Zealand Grassland Association 50: 181-187.
Hayes, B.J., N.O. Cogan, L.W. Pembleton, M.E. Goddard, J. Wang, G.C. Spangenberg, et al. 2013. Prospects for genomic selection in forage plant species. Plant Breeding 132: 133-143. doi:10.1111/pbr.12037.
Haynes, R. 1980. Competitive aspects of the grass-legume association. Advances in Agronomy 33: 227-261. doi:10.1016/S0065-2113(08)60168-6.
Hayward, M., N. McAdam, J. Jones, C. Evans, G. Evans, J. Forster, et al. 1994. Genetic markers and the selection of quantitative traits in forage grasses. Euphytica 77: 269-275. doi:10.1007/BF02262641.
Hedlund, K., I. Santa Regina, W. Van der Putten, J. Lepš, T. Diaz, G. Korthals, et al. 2003. Plant species diversity, plant biomass and responses of the soil community on abandoned land across Europe: idiosyncracy or above-belowground time lags. Oikos 103: 45-58. doi:10.1034/j.1600-0706.2003.12511.x.
Hedrick, P. 2011. Genetics of populationsJones & Bartlett Publishers, Boston, United States of America.
Helgadóttir, Á., S. Dalmannsdóttir and R.P. Collins. 2001. Adaptational changes in white clover populations selected under marginal conditions. Annals of Botany 88: 771-780. doi:10.1006/anbo.2001.1438,.

Herten, K., M.S. Hestand, J.R. Vermeesch and J.K. Van Houdt. 2015. GBSX: a toolkit for experimental design and demultiplexing genotyping by sequencing experiments. BMC Bioinformatics 16: 73. doi:10.1186/s12859-015-0514-3.

Hill Jr, R., J. Shenk and R. Barnes. 1988. Breeding for yield and quality. In: A. A. Hanson, D. K. Barnes and R. R. H. Jr., editors, Alfalfa and Alfalfa Improvement. p. 809-825.

Hill, W.G. 2014. Applications of Population Genetics to Animal Breeding, from Wright, Fisher and Lush to Genomic Prediction. Genetics 196: 1-16. doi:10.1534/genetics.112.147850.

Hill, W.G. and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38: 226-231. doi:10.1007/bf01245622.

Hirakawa, H., P. Kaur, K. Shirasawa, P. Nichols, S. Nagano, R. Appels, et al. 2016. Draft genome sequence of subterranean clover, a reference for genus *Trifolium*. Scientific Reports 6: 30358. doi:10.1038/srep30358.

Hofmann, R.W., E.E. Swinny, S.J. Bloor, K.R. Markham, K.G. Ryan, B.D. Campbell, et al. 2000. Responses of nine *Trifolium repens* L. populations to ultraviolet-B radiation: differential flavonol glycoside accumulation and biomass production. Annals of Botany 86: 527-537. doi:10.1006/anbo.2000.1216.

Hoglund, J., J. White, R. Burgess and J. Brock. 1985. Environmental and agronomic constraints in dryland pasture and choice of species. Using Herbage Cultivars. Grassland Research and Practice Series 3: 39-43.

Holland, J.B. 2007. Genetic architecture of complex traits in plants. Current Opinion in Plant Biology 10: 156-161. doi:10.1016/j.pbi.2007.01.003.

Holland, J.B., W.E. Nyquist and C.T. Cervantes-Martínez. 2003. Estimating and interpreting heritability for plant breeding: an update. Plant Breeding Reviews 22: 9-112. doi:10.1002/9780470650202.ch2.

Houston, K., S.M. McKim, J. Comadran, N. Bonar, I. Druka, N. Uzrek, et al. 2013. Variation in the interaction between alleles of HvAPETALA2 and microRNA172 determines the density of grains on the barley inflorescence. Proceedings of the National Academy of Sciences 110: 16675-16680. doi:10.1073/pnas.1311681110.

Hovin, A.W. 1962. Interspecific hybridization between *Trifolium repens* L. and *T. nigrescens* Viv. and analysis of hybrid meiosis Crop Science 2: 251-254.

Hoyos-Villegas, V., V. Arief, W.-H. Yang, M. Sun, I. DeLacy, B. Barrett, et al. 2018. QuLinePlus: Extending plant breeding strategy and genetic model simulation to cross-pollinated populations – case studies in forage breeding. Heredity. doi:10.5061/dryad.7368cc2.

Hoyos-Villegas, V., J. O'Connor, A. Heslop, A. Hilditch, M. Jahufer and B. Barrett. 2019. Rate of genetic gain for persistence to grazing and dry matter yield in white clover across 90 years of cultivar development. Crop Science. doi:10.2135/cropsci2018.07.0471.

Huang, J., J. Li, J. Zhou, L. Wang, S. Yang, L.D. Hurst, et al. 2018. Identifying a large number of high-yield genes in rice by pedigree analysis, whole-genome sequencing, and CRISPR-Cas9 gene knockout. Proceedings of the National Academy of Sciences 115: E7559-E7567. doi:10.1073/pnas.1806110115.

Huang, Y.-F., J.A. Poland, C.P. Wight, E.W. Jackson and N.A. Tinker. 2014. Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. PloS One 9. doi:10.1371/journal.pone.0102448.

Huber, H., E. Jacobs and E.J.W. Visser. 2008. Variation in flooding-induced morphological traits in natural populations of white clover (*Trifolium repens*) and their effects on plant performance during soil flooding. Annals of Botany 103: 377-386. doi:10.1093/aob/mcn149.

Hughes, M. 1991. The cyanogenic polymorphism in *Trifolium repens* L. (white clover). Heredity 66: 105-115.

Humphreys, M. 1997. The contribution of conventional plant breeding to forage crop improvement. Proceedings of the 18th International Grassland Congress'. Winnipeg and Saskatoon, Canada. p. 8-17.

Hunt, W. and H. Easton. 1989. Fifty years of ryegrass research in New Zealand. Proceedings of the New Zealand Grassland Association 50: 1-23.

Hussain, S., I. Verry, M. Jahufer and W. Williams. 2017. Cytological and morphological evaluation of interspecific hybrids between *Trifolium repens* and *T. uniflorum*. Crop Science 57: 2617-2625. doi:10.2135/cropsci2017.05.0314.

Hussain, S.W., I.M. Verry and W.M. Williams. 2016. Development of breeding populations from interspecific hybrids between *Trifolium repens* L. and *T. occidentale* Coombe. Plant Breeding 135: 118-123. doi:10.1111/pbr.12326.
Hussain, S.W. and W.M. Williams. 2013. *Trifolium occidentale*: a valuable genetic resource for white clover improvement. 22nd International Grasslands Congress: 309-310.
Hutchings, M.J., R. Turkington, E. Klein and P. Carey. 1997. Morphological plasticity in *Trifolium repens* L.: the effects of clone genotype, soil nutrient level, and the genotype of conspecific neighbours. Canadian Journal of Botany 75: 1382-1393. doi:10.1139/b97-852.
Hutchinson, K., K. King and D. Wilkinson. 1995. Effects of rainfall, moisture stress, and stocking rate on the persistence of white clover over 30 years. Australian Journal of Experimental Agriculture 35: 1039-1047. doi:10.1071/EA9951039.
Hyslop, M., P. Kemp and J. Hodgson. 1999. Vegetatively reproductive red clovers (*Trifolium pratense* L.): an overview. Proceedings of the Conference - New Zealand Grassland Association: 121-126.
Ikegawa, S. 2012. A short history of the genome-wide association study: where we were and where we are going. Genomics & Informatics 10: 220. doi:10.5808/GI.2012.10.4.220.
Inostroza, L., M. Bhakta, H. Acuña, C. Vásquez, J. Ibáñez, G. Tapia, et al. 2018. Understanding the complexity of cold tolerance in white clover using temperature gradient locations and a GWAS approach. The Plant Genome 11. doi:10.3835/plantgenome2017.11.0096.
IPCC. 2013. Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. In: T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, editors, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
Isobe, S., R. Kölliker, H. Hisano, S. Sasamoto, T. Wada, I. Klimenko, et al. 2009. Construction of a consensus linkage map for red clover (*Trifolium pratense* L.). BMC Plant Biology 9: 57.
Isobe, S., A. Sawai, H. Yamaguchi, M. Gau and K. Uchiyama. 2002. Breeding potential of the backcross progenies of a hybrid between *Trifolium medium* × *T. pratense* to *T. pratense*. Canadian Journal of Plant Science 82: 395-399. doi:10.4141/P01-034.
Ištvánek, J., M. Jaroš, A. Křenek and J. Řepková. 2014. Genome assembly and annotation for red clover (*Trifolium pratense*; Fabaceae). American Journal of Botany 101: 327-337. doi:10.3732/ajb.1300340.
Jahufer, M., M. Cooper, J. Ayres and R. Bray. 2002. Identification of research to improve the efficiency of breeding strategies for white clover in Australia - a review. Australian Journal of Agricultural Research 53: 239-257. doi:10.1071/AR01110.
Jahufer, M., M. Cooper, R. Bray and J. Ayres. 1999. Evaluation of white clover (*Trifolium repens* L.) populations for summer moisture stress adaptation in Australia. Australian Journal of Agricultural Research 50: 561-574.
Jahufer, M., A. Dunn, I. Baird, J. Ford, A. Griffiths, C. Jones, et al. 2013a. Genotypic variation for morphological traits in a white clover mapping population evaluated across two environments and three years. Crop Science 53: 460-472. doi:10.2135/cropsci2012.06.0370.
Jahufer, M., S. Nichols, J. Crush, L. Ouyang, A. Dunn, J. Ford, et al. 2008. Genotypic variation for root trait morphology in a white clover mapping population grown in sand. Crop Science 48: 487-494. doi:10.2135/cropsci2007.03.0161.
Jahufer, M.Z.Z., J.L. Ford, K.H. Widdup, C. Harris, G. Cousins, J.F. Ayres, et al. 2013b. Improving white clover for Australasia. Crop and Pasture Science 63: 739-745. doi:10.1071/CP12142.
Janick, J. 2003a. Plant Breeding Reviews. Wiley. doi:10.1002/9781118497869.
Jiang, Q., J.-Y. Zhang, X. Guo, M. Bedair, L. Sumner, J. Bouton, et al. 2010. Improvement of drought tolerance in white clover (*Trifolium repens*) by transgenic expression of a transcription factor gene WXP1. Functional Plant Biology 37: 157-165. doi:10.1071/FP09177.
Joensson, H. 1985. Red clover (*Trifolium pratense*) Sara. Agri Hortique Genetica.
Johnson, R. and T. Hodgkin. 1999a. Core collections for today and tomorrow. International Plant Genetic Resources Institute, Rome, Italy.
Johnson, R., W. Johnston, M. Nelson, C. Simon and C. Golob. 1999b. Core utilization and development–an example with *Poa pratensis*. Core Collections for Today and Tomorrow: 49-60.

Jones, E.S., L.J. Hughes, M.C. Drayton, M.T. Abberton, T.P. Michaelson-Yeates, C. Bowen, et al. 2003. An SSR and AFLP molecular marker-based genetic map of white clover (*Trifolium repens* L.). Plant Science 165: 531-539. doi:10.1016/S0168-9452(03)00212-7.
Jones, J.S. and E.T. Bingham. 2010. Inbreeding depression in alfalfa and cross-pollinated crops. Plant Breeding Reviews 13: 209-233. doi:10.1002/9780470650059.ch6.
Kahilainen, A., M. Puurtinen and J.S. Kotiaho. 2014. Conservation implications of species–genetic diversity correlations. Global Ecology and Conservation 2: 315-323. doi:10.1016/j.gecco.2014.10.013.
Kakes, P. 1997. Difference between the male and female components of fitness associated with the gene Ac in *Trifolium repens*. Acta Botanica Neerlandica 46: 219-223. doi:10.1111/plb.1997.46.2.219.
Kaler, A.S., J.D. Ray, W.T. Schapaugh, C.A. King and L.C. Purcell. 2017. Genome-wide association mapping of canopy wilting in diverse soybean genotypes. Theoretical and Applied Genetics 130: 2203-2217. doi:10.1007/s00122-017-2951-z.
Kannenberg, L. and F. Elliott. 1962. Ploidy in *Trifolium ambiguum* M. Bieb. in relation to some morphological and physiological characters. Crop Science 2: 378-381. doi:10.2135/cropsci1962.0011183X000200050004x.
Kasuga, M., S. Miura, K. Shinozaki and K. Yamaguchi-Shinozaki. 2004. A combination of the Arabidopsis DREB1A gene and stress-inducible rd29A promoter improved drought-and low-temperature stress tolerance in tobacco by gene transfer. Plant and Cell Physiology 45: 346-350. doi:10.1093/pcp/pch037.
Kaur, P., P.E. Bayer, Z. Milec, J. Vrána, Y. Yuan, R. Appels, et al. 2017. An advanced reference genome of *Trifolium subterraneum* L. reveals genes related to agronomic performance. Plant Biotechnology 15: 1034-1046. doi:10.1111/pbi.12697.
Kazimierska, E. 1978. Embryological studies of cross compatibility in the genus *Trifolium* L. II. Fertilization, development of embryo and endosperm in crossing *T. repens* L. with *T. medium* L. Plant Breeding Reviews 91: 51-24.
Keller, L.F. and D.M. Waller. 2002. Inbreeding effects in wild populations. Trends in Ecology & Evolution 17: 230-241. doi:10.1016/S0169-5347(02)02489-8.
Kelly, R., G. Shackell and A. Allison. 1980. Reproductive performance of ewes grazing red clover (Grasslands Pawera) or white clover—grass pasture at mating. New Zealand Journal of Experimental Agriculture 8: 87-91.
Kemp, P.D., C. Matthew and R.J. Lucas. 1999. Pasture species and cultivars. In: J. G. H. White and J. Hodgson, editors, New Zealand pasture and crop science. Oxford University Press, Oxford, United Kingdom. p. 83-89.
Kimura, M. and T. Ohta. 1971. Theoretical aspects of population genetics. Princeton University Press.
Knowles, I.M., T.J. Fraser and M.J. Daly. 2003. White clover: loss in drought and subsequent recovery. Legumes for dryland pastures. Grassland Research and Practice Series 11: 37-41.
Korte, A. and A. Farlow. 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9: 29. doi:10.1186/1746-4811-9-29.
Korte, A. and A.J. Parsons. 1984. Persistence of a large-leaved white clover variety under sheep grazing. Proceedings of the New Zealand Grassland Association.
Kouamé, C. and K. Quesenberry. 1993. Cluster analysis of a world collection of red clover germplasm. Genetic Resources and Crop Evolution 40: 39-47. doi:10.1007/BF00053463.
Kumar, S., T.W. Banks and S. Cloutier. 2012. SNP discovery through next-generation sequencing and its applications. International Journal of Plant Genomics 2012. doi:10.1155/2012/831460.
Kunert, K. and B.J. Vorster. 2020. In search for drought-tolerant soybean: is the slow-wilting phenotype more than just a curiosity? Journal of Experimental Botany 71: 457. doi:10.1093/jxb/erz235.
Lacape, M., J. Wery, D. Annerose and E. Jallas. 1998. Analysis of cotton genotypic differences for plant response to drought. Proceedings of the Beltwide Cotton Conference 2: 1383-1392.
Lacombe, T., J.-M. Boursiquot, V. Laucou, M. Di Vecchi-Staraz, J.-P. Péros and P. This. 2013. Large-scale parentage analysis in an extended set of grapevine cultivars (*Vitis vinifera* L.). Theoretical and Applied Genetics 126: 401-414. doi:10.1007/s00122-012-1988-2.
Lacy, R.C. 1989. Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. Zoo Biology 8: 111-123.
Ladizinsky, G. 1985. Founder effect in crop-plant evolution. Economic Botany 39: 191-199.

Lancashire, J. 2006. The importance of exotic germplasm to the NZ livestock industry. Breeding for success: diversity in action. Proceedings of the 13th Australasian Plant Breeding Conference: 1034-1041.
Laurens, F., M.J. Aranzana, P. Arus, D. Bassi, M. Bink, J. Bonany, et al. 2018. An integrated approach for increasing breeding efficiency in apple and peach in Europe. Horticulture Research 5: 1-14. doi:10.1038/s41438-018-0016-3.
Lecoeur, J. and T.R. Sinclair. 1996. Field pea transpiration and leaf growth in response to soil water deficits. Crop Science 36: 331-335. doi:10.2135/cropsci1996.0011183X003600020020x.
Ledgard, S., G. Brier and M. Upsdel. 1990. Effect of clover cultivar on production and nitrogen fixation in clover-ryegrass swards under dairy cow grazing. New Zealand Journal of Agricultural Research 33: 243-249. doi:10.1080/00288233.1990.10428416.
Lee, C., H. Eagles, J. Caradus and K. Reed. 1993. Investigation of yield and persistence of white clover using cluster analyses. Euphytica 72: 219-224.
Leroy, G., X. Rognon, A. Varlet, C. Joffrin and E. Verrier. 2006. Genetic variability in French dog breeds assessed by pedigree data. Journal of Animal Breeding and Genetics 123: 1-9. doi:10.1111/j.1439-0388.2006.00565.x.
Levy, E.B. 1970. Grasslands of New Zealand. Wellington, A.R. Shearer, Government Printer.
Lewis, G.P. 2005. Legumes of the World. Royal Botanic Gardens.
Lewontin, R.C. 1964. The Interaction of Selection and Linkage. I. General Considerations; Heterotic Models. Genetics 49: 49-67.
Li, H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987-2993. doi:10.1093/bioinformatics/btr509.
Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler Transform. Bioinformatics 25: 1754-1760. doi:10.1093/bioinformatics/btp324.
Li, L., Q. Zhang and D. Huang. 2014a. A review of imaging techniques for plant phenotyping. Sensors 14: 20078-20111. doi:10.3390/s141120078.
Li, X., Y. Han, Y. Wei, A. Acharya, A.D. Farmer, J. Ho, et al. 2014b. Development of an alfalfa SNP array and its use to evaluate patterns of population structure and linkage disequilibrium. PloS One 9. doi:10.1371/journal.pone.0084329.
Li, X., Y. Wei, A. Acharya, J.L. Hansen, J.L. Crawford, D.R. Viands, et al. 2015. Genomic prediction of biomass yield in two selection cycles of a tetraploid alfalfa breeding population. The Plant Genome 8. doi:10.3835/plantgenome2014.12.0090.
Li, Z., Y. Peng and X. Ma. 2013a. Different response on drought tolerance and post-drought recovery between the small-leafed and the large-leafed white clover (*Trifolium repens* L.) associated with antioxidative enzyme protection and lignin metabolism. Acta Physiologiae Plantarum 35: 213-222. doi:10.1007/s11738-012-1066-z.
Li, Z., Y. Peng, X.-Q. Zhang, M.-H. Pan, X. Ma, L.-K. Huang, et al. 2014c. Exogenous spermidine improves water stress tolerance of white clover (*Trifolium repens* L.) involved in antioxidant defence, gene expression and proline metabolism. Plant Omics 7: 517-526.
Li, Z., P. Shi and Y. Peng. 2013b. Improved drought tolerance through drought preconditioning associated with changes in antioxidant enzyme activities, gene expression and osmoregulatory solutes accumulation in white clover (*Trifolium repens* L.). Plant Omics 6: 481-489.
Liatukas, Ž. and J. Bukauskaitė. 2012. Differences in yield of diploid and tetraploid red clover in Lithuania. Proceedings of the Latvian Academy of Sciences 66: 163-167. doi:10.2478/v10046-012-0023-y.
Liu, J., Z. Wang, Y. Wang, R. Li and R. Wu. 2012. Model and algorithm for linkage disequilibrium analysis in a non-equilibrium population. Frontiers in Genetics 3: 78. doi:10.3389/fgene.2012.00078.
Liu, J., C. Yang, X. Shi, C. Li, J. Huang, H. Zhao, et al. 2016. Analyzing association mapping in pedigree-based GWAS using a penalized multitrait mixed model. Genetic Epidemiology 40: 382-393. doi:10.1002/gepi.21975.
Liu, R., J. Gong, X. Xiao, Z. Zhang, J. Li, A. Liu, et al. 2018. GWAS analysis and QTL identification of fiber quality traits and yield components in upland cotton using enriched high-density SNP markers. Frontiers in Plant Science 9: 1067. doi:10.3389/fpls.2018.01067.

Liu, X.-P. and L.-X. Yu. 2017. Genome-wide association mapping of loci associated with plant growth and forage production under salt stress in alfalfa (*Medicago sativa* L.). Frontiers in Plant Science 8: 853. doi:10.3389/fpls.2017.00853.
López, A., J. Maiztegui and R. Cabrera. 1998. Voluntary intake and digestibility of forages with different nutritional quality in alpacas (*Lama pacos*). Small Ruminant Research 29: 295-301. doi:10.1016/S0921-4488(97)00135-1.
Lotscher, M. and M. Hay. 1997. Genotypic differences in physiological integration, morphological plasticity and utilization of phosphorus induced by variation in phosphate supply in *Trifolium repens*. Journal of Ecology: 341-350. doi:10.2307/2960506
Lucas, R.J., A. Mills, S. Wright, A. Black and D.J. Moot. 2015. Selection of sub clover cultivars for New Zealand dryland pastures. Journal of New Zealand Grasslands 77: 203-210.
Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Sunderland, MA.
Ma, X., F. Feng, H. Wei, H. Mei, K. Xu, S. Chen, et al. 2016. Genome-wide association study for plant height and grain yield in rice under contrasting moisture regimes. Frontiers in Plant Science 7: 1801. doi:10.3389/fpls.2016.01801.
Ma, X., Z. Wang, W. Li, Y. Zhang, X. Zhou, Y. Liu, et al. 2019. Resequencing core accessions of a pedigree identifies derivation of genomic segments and key agronomic trait loci during cotton improvement. Plant Biotechnology Journal 17: 762-775. doi:10.1111/pbi.13013.
Macfarlane, M., A. McGowan, G. Sheath and C. Korte. 1990a. An on-farm evaluation of white and subterranean clovers in North Island hill country. Proceedings of the New Zealand Grassland Association 51: 157-161.
Macfarlane, M. and G. Sheath. 1984. Clover - what types for dry hill country? Proceedings of the New Zealand Grassland Association 45: 140-150.
Macfarlane, M., G. Sheath and M. Tucker. 1990b. Evaluation of clovers in dry hill country 6. Subterranean and white clovers at Wairakei, New Zealand. New Zealand Journal of Agricultural Research 33: 557-564. doi:10.1080/00288233.1990.10428457.
Mackay, A. 1991. Performance of white clover cultivars and breeding lines in a mixed species sward: 1. Yield and clover content. New Zealand Journal of Agricultural Research 34: 141-154. doi:10.1080/00288233.1991.10423353.
Mackay, G.R. 1973. Interspecific hybrids between forage rape (*Brassica napus* L.) and turnip (*Brassica campestris* L. SSP. *Rapifera*) as alternatives to forage rape. 1. An exploratory study with single pair crosses. Euphytica 22: 495-499. doi:10.1007/BF00036646.
Mackay, T.F. 2001. The genetic architecture of quantitative traits. Annual Review of Genetics 35: 303-339. doi:10.1146/annurev.genet.35.102401.090633.
Macleod, C., M.W. Humphreys, W.R. Whalley, L. Turner, A. Binley, C.W. Watts, et al. 2013. A novel grass hybrid to reduce flood generation in temperate regions. Scientific Reports 3: 1-7. doi:10.1038/srep01683.
Majack, W., T.A. McAllister, D. McCartney, K. Stanford and K.-J. Cheng. 2003. Bloat in cattle. Alberta Agriculture, Food and Rural Development, Information Packing Centre.
Malaviya, D., A. Roy, P. Kaushal, B. Kumar, A. Tiwari and C. Lorenzoni. 2004. Development and characterization of interspecific hybrids of *Trifolium alexandrinum* × *T. apertum* using embryo rescue. Plant Breeding 123: 536-542. doi:10.1111/j.1439-0523.2004.01042.x.
Malaviya, D.R., A.K. Roy, P. Kaushal, M. Chakraborti, A. Yadav, A. Khare, et al. 2018. Interspecific compatibility barriers, development of interspecific hybrids through embryo rescue and lineage of *Trifolium alexandrinum* (Egyptian clover)—important tropical forage legume. Plant Breeding 137: 655-672. doi:10.1111/pbr.12616.
Manolio, T.A., F.S. Collins, N.J. Cox, D.B. Goldstein, L.A. Hindorff, D.J. Hunter, et al. 2009. Finding the missing heritability of complex diseases. Nature 461: 747. doi:10.1038/nature08494.
Marita, J.M., J.M. Rodriguez and J. Nienhuis. 2000. Development of an algorithm identifying maximally diverse core collections. Genetic Resources and Crop Evolution 47: 515-526. doi:10.1023/A:1008784610962.
Marshall, A., T. Michaelson-Yeates, P. Aluka and M. Meredith. 1995. Reproductive characters of interspecific hybrids between *Trifolium repens* L. and *Trifolium nigrescens* Viv. Heredity 74: 136-145. doi:10.1038/hdy.1995.20.
Marshall, A., T.P. Michaelson-Yeates and M. Abberton. 2008. Introgression of reproductive traits from *Trifolium nigrescens* increases the seed yield of white clover (*T. repens*). Plant Breeding 127: 597-601. doi:10.1111/j.1439-0523.2008.01534.x.

Marshall, A.H., M. Lowe and R.P. Collins. 2015. Variation in response to moisture stress of young plants of interspecific hybrids between white clover (*T. repens* L.) and Caucasian clover (*T. ambiguum* M. Bieb.). Agriculture 5: 353-366. doi:10.3390/agriculture5020353.

Marshall, A.H., C. Rascle, M.T. Abberton, T.P.T. Michaelson-Yeates and I. Rhodes. 2001. Introgression as a route to improved drought tolerance in white clover (*Trifolium repens* L.). Journal of Agronomy and Crop Science 187: 11-18. doi:10.1046/j.1439-037X.2001.00495.x.

Marshall, A.H., A. Williams, M.T. Abberton, T.P. Michaelson-Yeates and H.G. Powell. 2003. Dry matter production of white clover (*Trifolium repens* L.), Caucasian clover (*T. ambiguum* M. Bieb.) and their associated hybrids when grown with a grass companion over 3 harvest years. Grass and Forage Science 58: 63-69. doi:10.1046/j.1365-2494.2003.00354.x.

Marshall, A.H., T.A. Williams, M.T. Abberton, T.P. Michaelson-Yeates, P. Olyott and H.G. Powell. 2004. Forage quality of white clover (*Trifolium repens* L.) × Caucasian clover (*T. ambiguum* M. Bieb.) hybrids and their grass companion when grown over three harvest years. Grass and Forage Science 59: 91-99. doi:10.1111/j.1365-2494.2004.00409.x.

Mather, R., D. Melhuish and M. Herlihy. 1996. Trends in the global marketing of white clover cultivars. Special Publication - Agronomy Society of New Zealand: 7-14.

Matsumoto, Y., T. Goto, J. Nishino, H. Nakaoka, A. Tanave, T. Takano-Shimizu, et al. 2017. Selective breeding and selection mapping using a novel wild-derived heterogeneous stock of mice revealed two closely-linked loci for tameness. Scientific Reports 7: 1-14. doi:10.1038/s41598-017-04869-1.

Matthews, D.L. and W.R. Battle. 1951. A survey of variability in alsike clover (*Trifolium hybridum* L.). Agronomy Journal 43: 45-46. doi:10.2134/agronj1951.00021962004300010010x.

Maxted, N. and S.J. Bennett. 2001. Plant Genetic Resources of Legumes in the Mediterranean. 39. doi:10.1007/978-94-015-9823-1.

McCoy, T. 1985. Interspecific hybridization of *Medicago sativa* L. and *M. rupestris* MB using ovule-embryo culture. Canadian Journal of Genetics and Cytology 27: 238-245. doi:10.1139/g85-035.

McCoy, T. and L. Smith. 1984. Uneven ploidy levels and a reproductive mutant required for interspecific hybridization of *Medicago sativa* L. × *Medicago dzhawakhetica* Bordz. Canadian Journal of Genetics and Cytology 26: 511-518. doi:10.1139/g84-081.

McCoy, T. and L. Smith. 1986. Interspecific hybridization of perennial *Medicago* species using ovule-embryo culture. Theoretical and Applied Genetics 71: 772-783. doi:10.1007/BF00276417.

McNally, K.L., K.L. Childs, R. Bohnert, R.M. Davidson, K. Zhao, V.J. Ulat, et al. 2009. Genome-wide SNP variation reveals relationships among landraces and modern varieties of rice. Proceedings of the National Academy of Sciences 106: 12273-12278.

Melchinger, A.E., A. Graner, M. Singh and M.M. Messmer. 1994. Relationships among European barley germplasm: I. Genetic diversity among winter and spring cultivars revealed by RFLPs. Crop Science 34: 1191-1199. doi:10.2135/cropsci1994.0011183X003400050009x.

Mercer, C., J. Van Den Bosch and K. Miller. 2000. Progress in recurrent selection and in crossing cultivars with white clover resistant to the clover rootknot nematode *Meloidogyne trifoliophila*. New Zealand Journal of Agricultural Research 43: 41-48. doi:10.1080/00288233.2000.9513407.

Mercer, C.F., J. Van Den Bosch and K.J. Miller. 1999. Effectiveness of recurrent selection of white clover (*Trifolium repens*) for resistance to New Zealand populations of clover cyst nematode (*Heterodera trifolii*). Nematology 1: 449-455. doi:10.1163/156854199508432.

Meredith, M.R., T.P.T. Michaelson-Yeates, H.J. Ougham and H. Thomas. 1995. *Trifolium ambiguum* as a source of variation in the breeding of white clover. Euphytica 82: 185-191. doi:10.1007/BF00027065.

Merker, A. 1984. Hybrids between *Trifolium medium* and *Trifolium pratense*. Hereditas 101: 267-268. doi:10.1111/j.1601-5223.1984.tb00927.x.

Meyer, W. and G. Green. 1981. Plant indicators of wheat and soybean crop water stress. Irrigation Science 2: 167-176. doi:10.1007/BF00257978.

Michaelson-Yeates, T.P.T., A. Marshall, M.T. Abberton and I. Rhodes. 1997. Self-compatibility and heterosis in white clover (*Trifolium repens L.*). Euphytica 94: 341-348. doi:10.1023/a:1002989410326.

Miller, G.L. 2000. Physiological response of bermudagrass grown in soil amendments during drought stress. Hort Science 35: 213-216. doi:10.21273/HORTSCI.35.2.213.

Minitab, L. 2006. Minitab. Inc., version 15.
Moot, D., W. Scott, A. Roy and A. Nicholls. 2000. Base temperature and thermal time requirements for germination and emergence of temperate pasture species. New Zealand Journal of Agricultural Research 43: 15-25.
Morley, F., A. Axelsen and D. Bennett. 1964. Effects of grazing red clover (*Trifolium pratense* L.) during the joining season on ewe fertility. Proceedings of the Australian Society of Animal Production 5: 58-61.
Morris, J. and S. Greene. 2001. Defining a multiple-use germplasm collection for the genus *Trifolium*. Crop Science 41: 893-901. doi:10.2135/cropsci2001.413893x.
Mosjidis, J.A. and K.A. Klingler. 2006. Genetic diversity in the core subset of the U.S. red clover germplasm. Crop Science 46: 758-762. doi:10.2135/cropsci2005.05-0076.
Muchow, R. and T. Sinclair. 1991. Water deficit effects on maize yields modeled under current and "greenhouse" climates. Agronomy Journal 83: 1052-1059. doi:10.2134/agronj1991.00021962008300060023x.
Mullan, B., A. Porteous, D. Wratt and M. Hollis. 2005. Changes in drought risk with climate change. Prepared for Ministry for the Environment (NZ Climate Change Office) and Ministry of Agriculture and Forestry. NIWA Client Report: WLG2005–23.(National Institute of Water and Atmospheric Research, Wellington).
Nadolska-Orczyk, A., I.K. Rajchel, W. Orczyk and S. Gasparis. 2017. Major genes determining yield-related traits in wheat and barley. Theoretical and Applied Genetics 130: 1081-1098. doi:10.1007/s00122-017-2880-x.
Nass, L., M. Sigrist, C. Ribeiro and F. Reifschneider. 2012. Genetic resources: the basis for sustainable and competitive plant breeding. Crop Breeding and Applied Biotechnology 12: 75-86. doi:10.1590/S1984-70332012000500009.
Nass, L.L. and E. Paterniani. 2000. Pre-breeding: a link between genetic resources and maize breeding. Scientia Agricola 57: 581-587. doi:10.1590/S0103-90162000000300035.
Navabi, A., P. Balasubramanian, K. Pauls, K. Bett and A. Hou. 2014. Genetic diversity of the Canadian dry bean varieties released since 1930: A pedigree analysis. Crop Science 54: 993-1003. doi:10.2135/cropsci2013.04.0210.
Nenz, E., F. Pupilli, F. Damiani and S.t. Arcioni. 1996. Somatic hybrid plants between the forage legumes *Medicago sativa* L. and *Medicago arborea* L. Theoretical and Applied Genetics 93: 183-189. doi:10.1007/BF00225744.
Nicholas, D., D. Gillespie, J. Smith and V. Hays. 1981. Procedure for selecting subterranean clover cultivars in South Western Australia. Proceedings of the XIV International Grassland Congress: 135-137.
Nichols, P., K. Foster, E. Piano, L. Pecetti, P. Kaur, K. Ghamkhar, et al. 2013a. Genetic improvement of subterranean clover (*Trifolium subterraneum* L.). 1. Germplasm, traits and future prospects. Crop and Pasture Science 64: 312-346. doi:10.1071/CP13118.
Nichols, P., R. Jones, T. Ridsdill-Smith and M. Barbetti. 2014a. Genetic improvement of subterranean clover (*Trifolium subterraneum* L.). 2. Breeding for disease and pest resistance. Crop and Pasture Science 65: 1207-1229. doi:10.1071/CP14031.
Nichols, P., A. Loi, B. Nutt, P. Evans, A. Craig, B. Pengelly, et al. 2007a. New annual and short-lived perennial pasture legumes for Australian agriculture - 15 years of revolution. Field Crops Research 104: 10-23. doi:10.1016/j.fcr.2007.03.016.
Nichols, P., C. Revell, A. Humphries, J. Howie, E. Hall, G. Sandral, et al. 2013b. Temperate pasture legumes in Australia - their history, current use, and future prospects. Crop and Pasture Science 63: 691-725. doi:10.1071/CP12194.
Nichols, S., J. Crush and L. Ouyang. 2014b. Phosphate responses of some *Trifolium repens* × *T. uniflorum* interspecific hybrids grown in soil. Crop and Pasture Science 65: 382-387. doi:10.1071/CP14029.
Nichols, S., R. Hofmann and W. Williams. 2014c. Drought resistance of *Trifolium repens* ×*Trifolium uniflorum* interspecific hybrids. Crop and Pasture Science 65: 911-921. doi:10.1071/CP14067.
Nichols, S., R. Hofmann and W. Williams. 2014d. The effect of interspecific hybridisation with *Trifolium uniflorum* on key white clover characteristics. Field Crops Research 161: 107-117. doi:10.1016/j.fcr.2014.03.004.

Nichols, S., R. Hofmann, W. Williams and J. Crush. 2014e. Nutrient responses and macronutrient composition of some *Trifolium repens* × *Trifolium uniflorum* interspecific hybrids. Crop and Pasture Science 65: 370-381. doi:10.1071/CP13446.
Nichols, S., R. Hofmann, W. Williams and C. van Koten. 2016. Rooting depth and root depth distribution of *Trifolium repens* × *T. uniflorum* interspecific hybrids. Annals of Botany 118: 699-710. doi:10.1093/aob/mcw067.
Nichols, S.N., J.R. Crush and D.R. Woodfield. 2007b. Effects of inbreeding on nodal root system morphology and architecture of white clover (*Trifolium repens* L.). Euphytica 156: 365-373. doi:10.1007/s10681-007-9386-6.
Nichols, S.N., R. Hofmann, I.M. Verry and W.M. Williams. 2013c. Improved drought stress tolerance of white clover through hybridisation with *Trifolium uniflorum* L. Proceedings of the 22nd International Grassland Congress.
Nichols, S.N., R.W. Hofmann and W.M. Williams. 2015. Physiological drought resistance and accumulation of leaf phenolics in white clover interspecific hybrids. Environmental and Experimental Botany 119: 40-47. doi:10.1016/j.envexpbot.2015.05.014.
Nikovitz, A. 1985. Bee pasture value of red clover varieties. Meheszet 33: 7.
NIWA. 2017. New Zealand Climate Update.
Novo, P.E., J.F.M. Valls, F. Galdeano, A.I. Honfi, F. Espinoza and C.L. Quarin. 2016. Interspecific hybrids between *Paspalum plicatulum* and *P. oteroi*: a key tool for forage breeding. Scientia Agricola 73: 356-362. doi:10.1590/0103-9016-2015-0218
Nowosad, F. and R. MacVicar. 1940. Adaptation of the "picric-acid test" method for selecting HCN-free lines in Sudan grass. Scientific Agriculture 20: 566-569. doi:10.4141/sa-1940-0038.
Nutman, P. and J. Riley. 1981. Breeding of nodulated red clover (*Trifolium pratense*) for high yield. Annals of Applied Biology 98: 319-331. doi:10.1111/j.1744-7348.1981.tb00764.x.
Olsen, K., B. Sutherland and L. Small. 2007. Molecular evolution of the Li/li chemical defence polymorphism in white clover (*Trifolium repens* L.). Molecular Ecology 16: 4180-4193. doi:10.1111/j.1365-294X.2007.03506.x.
Olykan, S.T., R.J. Lucas, C.S. Teixeira, R.A. Subtil and D.J. Moot. 2018. Establishment, production and regeneration of subterranean clovers in the Mackenzie Basin, New Zealand. Journal of New Zealand Grasslands.
Oram, R.N. 1977. *Trifolium ambiguum* M. Bieb. (Caucasian clover) cv. Monaro. Journal of the Australian Institute of Agricultural Science 43: 155-161.
Orr, R., A. Parsons, P. Penning and T. Treacher. 1990. Sward composition, animal performance and the potential production of grass/white clover swards continuously stocked with sheep. Grass and Forage Science 45: 325-336. doi:10.1111/j.1365-2494.1990.tb01957.x.
Ortega, F., L. Parra and A. Quiroz. 2014. Breeding red clover for improved persistence in Chile: a review. Crop and Pasture Science 65: 1138-1146. doi:10.1071/CP13323.
Paiva, S.R., O. Facó, D.A. Faria, T. Lacerda, G.B. Barretto, P.L. Carneiro, et al. 2011. Molecular and pedigree analysis applied to conservation of animal genetic resources: the case of Brazilian Somali hair sheep. Tropical Animal Health and Production 43: 1449-1457. doi:10.1007/s11250-011-9873-6.
Palmer, T. 1972. Variation in flowering time among and within populations of *Trifolium arvense* L. in New Zealand. New Zealand Journal of Botany 10: 59-68. doi:10.1080/0028825X.1972.10430211.
Pantalone, V., G. Rebetzke, J. Burton and T. Carter Jr. 1996. Phenotypic evaluation of root traits in soybean and applicability to plant breeding. Crop Science 36: 456-459. doi:10.2135/cropsci1996.0011183X003600020039x.
Patisaul, H.B. and W. Jefferson. 2010. The pros and cons of phytoestrogens. Frontiers in neuroendocrinology 31: 400-419. doi:10.1016/j.yfrne.2010.03.003.
Pecetti, L., P. Annicchiarico, A. Abdelguerfi, R. Kallida, M. Mefti, C. Porqueddu, et al. 2011. Response of Mediterranean tall fescue cultivars to contrasting agricultural environments and implications for selection. Journal of Agronomy and Crop Science 197: 12-20. doi:10.1111/j.1439-037X.2010.00443.x.
Pederson, G.A. and G.E. Brink. 1998. Cyanogenesis effect on insect damage to seedling white clover in a bermudagrass sod. Agronomy Journal 90: 208-210. doi:10.2134/agronj1998.00021962009000020015x.

Pederson, G.A. and M.R. McLaughlin. 1989. Resistance to viruses in *Trifolium* interspecific hybrids related to white clover. Plant Disease 73: 997-999.
Peeters, J. and J. Williams. 1984. Towards better use of genebanks with special reference to information. Plant Genetic Resource Newsletter 60: 22-32.
Pham, A.-T., A. Maurer, K. Pillen, C. Brien, K. Dowling, B. Berger, et al. 2019. Genome-wide association of barley plant growth under drought stress using a nested association mapping population. BMC Plant Biology 19: 134. doi:10.1186/s12870-019-1723-0.
Philipp, N., S. Weise, M. Oppermann, A. Börner, A. Graner, J. Keilwagen, et al. 2018. Leveraging the use of historical data gathered during seed regeneration of an ex situ genebank collection of wheat. Frontiers in Plant Science 9: 609. doi:10.3389/fpls.2018.00609.
Phosphate, I., R.W. Snaydon and A.D. Bradshaw. 1962. Differences between natural populations of *Trifolium repens* L. in response to mineral nutrients. Journal of Experimental Botany 13: 422-434. doi:10.1093/jxb/13.3.422.
Plucknett, D., N. Smith, J. Willaims and N. Anishetty. 1987. Gene banks and the world's food. Princeton University Press.
Priolli, R.H.G., J.B. Pinheiro, M.I. Zucchi, M.M. Bajay and N.A. Vello. 2010. Genetic diversity among Brazilian soybean cultivars based on SSR loci and pedigree data. Brazilian Archives of Biology and Technology 53: 519-531. doi:10.1590/S1516-89132010000300004
Pritchard, H., K. Manger and F. Prendergast. 1988. Changes in *Trifolium arvense* seed quality following alternating temperature treatment using liquid nitrogen. Annals of Botany 62: 1-11. doi:10.1093/oxfordjournals.aob.a087626.
Pungulani, L.L.M. 2014. Exploring the genetic potential of locally adapted germplasm for drought tolerance: a case for cowpea (*Vigna unguiculata* (L.)) Walp) from Malawi. Doctoral, Massey University.
Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. The American Journal of Human Genetics 81: 559-575. doi:10.1086/519795.
Quesenberry, K., G. Prine, O. Ruelke, L. Dunavin and P. Mislevy. 1993. Registration of 'Cherokee 'red clover. Crop Science 33: 208-209. doi:10.2135/cropsci1993.0011183X003300010051x.
Quesenberry, K. and N. Taylor. 1977. Interspecific in *Trifolium* L. Sect. *Trifolium* Zoh. II. Fertile Ploypoid Hybrids Between *T. medium* L. and *T. sarosiense* Hazsl. Crop Science 17: 141-145. doi:10.2135/cropsci1977.0011183X001700010037x.
Raj, A., M. Stephens and J.K. Pritchard. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics 197: 573-589. doi:10.1534/genetics.114.164350.
Rambaut, A. 2009. FigTree, version 1.3. 1. Computer program distributed by the author.
Ramírez-Villegas, J., C. Khoury, A. Jarvis, D.G. Debouck and L. Guarino. 2010. A gap analysis methodology for collecting crop genepools: a case study with *Phaseolus* beans. PloS One 5: e13497. doi:10.1371/journal.pone.0013497.
Rattray, P.V. 2005. Clover management, research, development & extension in the New Zealand pastoral industries. In: M. o. P. Industry, editor Wellington, New Zealand.
Ray, J.D. and T.R. Sinclair. 1997. Stomatal closure of maize hybrids in response to drying soil. Crop Science 37: 803-807. doi:10.2135/cropsci1997.0011183X003700030018x.
Ray, J.D. and T.R. Sinclair. 1998. The effect of pot size on growth and transpiration of maize and soybean during water deficit stress. Journal of Experimental Botany 49: 1381-1386. doi:10.1093/jxb/49.325.1381.
Reed, D.H. and R. Frankham. 2003. Correlation between fitness and genetic diversity. Conservation Biology 17: 230-237. doi:10.1046/j.1523-1739.2003.01236.x.
Reynolds, R.G. and B.M. Fitzpatrick. 2013. Tests of two methods for identifying founder effects in metapopulations reveal substantial type II error. Genetica 141: 119-131. doi:10.1007/s10709-013-9711-z.
Rhodes, I. and W. Harris. 1979. Nature and basis of differences in sward composition and yield in ryegrass-white clover mixtures. Occasional Symposium.
Rhodes, I. and K.J. Webb. 1993. Improvement of white clover. Outlook on Agriculture 22: 189-194. doi:10.1177/003072709302200310.
Richards, C.M. and G.M. Volk. 2010. New Challenges for Data Management in Genebanks. Acta Horticulturae 859: 333-336. doi:10.17660/ActaHortic.2010.859.39.

Riday, H. 2010. Progress made in improving red clover (*Trifolium pratense* L.) through breeding. International Journal of Plant Breeding 4: 22-29.
Riday, H. 2011. Paternity testing: a non-linkage based marker-assisted selection scheme for outbred forage species. Crop Science 51: 631-641. doi:10.2135/cropsci2010.07.0390.
Ritchie, J.T. 1973. Influence of soil water status and meteorological conditions on evaporation from a corn canopy. Agronomy Journal 65: 893-897. doi:10.2134/agronj1973.00021962006500060014x.
Robinson, G. and A. Lazenby. 1976. Effect of superphosphate, white clover and stocking rate on the productivity of natural pastures, Northern Tablelands, New South Wales. Australian Journal of Experimental Agriculture 16: 209-217. doi:10.1071/EA9760209.
Roldán-Ruiz, I. and R. Kölliker. 2010. Marker-assisted selection in forage crops and turf: a review. In: C. Huyghe, editor Sustainable use of genetic diversity in forage and turf breeding. Springer. p. 383-390.
Rosenthal, W., G. Arkin, P. Shouse and W. Jordan. 1987. Water deficit effects on transpiration and leaf growth. Agronomy Journal 79: 1019-1026. doi:10.2134/agronj1987.00021962007900060014x.
Rosso, B. and E. Pagano. 2005. Evaluation of introduced and naturalised populations of red clover (*Trifolium pratense* L.) at Pergamino EEA-INTA, Argentina. Genetic Resources and Crop Evolution 52: 507-511. doi:10.1007/s10722-005-0777-z.
Rosso, B.S. and E.M. Pagano. 2001. Collection and characterization of naturalized populations of white clover (*Trifolium repens* L.) in Argentina. Genetic Resources and Crop Evolution 48: 513-517. doi:10.1023/A:1012005800481.
Rotili, P. 1991. The role of selfing in the synthetic and hybrid varieties constitution in forage crops. Proceedings of the 16th Meeting of Fodder Crops Breeding of Eucarpia: 83-88.
Roughsedge, T., S. Brotherstone and P. Visscher. 1999. Quantifying genetic contributions to a dairy cattle population using pedigree analysis. Livestock Production Science 60: 359-369. doi:10.1016/S0301-6226(99)00106-2.
Rubenstein, K.D., M. Smale and M.P. Widrlechner. 2006. Demand for genetic resources and the US National Plant Germplasm System. Crop Science 46: 1021-1031. doi:10.2135/cropsci2005.0129.
Rumball, W. and C. Armstrong. 1974. The performance of overseas ryegrass cultivars in New Zealand. Proceedings of the New Zealand Grassland Association 36: 97-104.
Rumball, W. and R. Claydon. 2005. 'G41 'Zigzag clover (*Trifolium medium* L.). New Zealand Journal of Agricultural Research. doi:10.1080/00288233.2005.9513642.
Rumball, W., R. Keogh, J. Miller and R. Claydon. 1997. 'Grasslands G27' red clover (*Trifolium pratense* L.). New Zealand Journal of Agricultural Research 40: 369-372.
Russel, J. and H. Webb. 1976. Climatic range of grasses and legumes used in pastures. Australian Institute of Agricultural Science 42: 156-166.
Sadok, W. and T.R. Sinclair. 2010. Transpiration response of 'slow-wilting'and commercial soybean (*Glycine max* (L.) Merr.) genotypes to three aquaporin inhibitors. Journal of Experimental Botany 61: 821-829. doi:10.1093/jxb/erp350.
Saeidnia, F., M.M. Majidi, A. Mirlohi and B. Ahmadi. 2018. Physiological responses of drought tolerance in orchardgrass (*Dactylis glomerata*) in association with persistence and summer dormancy. Crop and Pasture Science 69: 515-526. doi:10.1071/CP17314.
Saeidnia, F., M.M. Majidi, A. Mirlohi and S. Shahidaval. 2016. Selection for productivity, persistence and drought tolerance in orchardgrass. Euphytica 212: 111-130. doi:10.1007/s10681-016-1776-1.
Sakiroglu, M. and E.C. Brummer. 2017. Identification of loci controlling forage yield and nutritive value in diploid alfalfa using GBS-GWAS. Theoretical and Applied Genetics 130: 261-268. doi:10.1007/s00122-016-2782-3.
Sakiroglu, M., S. Sherman-Broyles, A. Story, K.J. Moore, J.J. Doyle and E.C. Brummer. 2012. Patterns of linkage disequilibrium and association mapping in diploid alfalfa (*M. sativa* L.). Theoretical and Applied Genetics 125: 577-590. doi:10.1007/s00122-012-1854-2.
Sanderson, M., R. Byers, R. Skinner and G. Elwinger. 2003. Growth and complexity of white clover stolons in response to biotic and abiotic stress. Crop Science 43: 2197-2205. doi:10.2135/cropsci2003.2197.

Sanderson, M., R. Skinner, D. Barker, G. Edwards, B. Tracy and D. Wedin. 2004. Plant species diversity and management of temperate forage and grazing land ecosystems. Crop Science 44: 1132-1144. doi:10.2135/cropsci2004.1132.

Sato, S., S. Isobe, E. Asamizu, N. Ohmido, R. Kataoka, Y. Nakamura, et al. 2005. Comprehensive structural analysis of the genome of red clover (*Trifolium pratense* L.). DNA research 12: 301-364. doi:10.1093/dnares/dsi018.

Sawai, A., S. Ueda, M. Gau and K. Uchiyama. 1990. Interspecific hybrids of *Trifolium medium* L. × 4x *T. pratense* L. obtained through embryo culture. Japanese Journal of Grassland Science 35: 267-272. doi:10.14941/grass.35.267.

Sawai, A., H. Yamaguchi and K. Uchiyama. 1995. Fertility and morphology of the chromosome-doubled hybrid *Trifolium medium* × *T. pratense* (red clover) and backcross progeny. Japanese Journal of Grassland Science 41: 122-127.

Scoppola, A., J.L. Tirado, F.M. Gutiérrez and S. Magrini. 2018. The genus *Trifolium* (*Fabaceae*) in south Europe: a critical review on species richness and distribution. Nordic Journal of Botany 36. doi:10.1111/njb.01723.

Sebastiani, P., N. Timofeev, D.A. Dworkis, T.T. Perls and M.H. Steinberg. 2009. Genome-wide association studies and the genetic dissection of complex traits. American Journal of Hematology 84: 504-515. doi:10.1002/ajh.21440.

Sehgal, D., P. Vikram, C.P. Sansaloni, C. Ortiz, C. Saint Pierre, T. Payne, et al. 2015. Exploring and mobilizing the gene bank biodiversity for wheat improvement. PloS One 10: e0132112. doi:10.1371/journal.pone.0132112.

Shackell, G., R. Kelly and P. Johnstone. 1993a. Effects of prolonged exposure of ewes to oestrogenic pasture 1. Permanent flock infertility following long-term grazing of red clover ('Grasslands Pawera')-dominant pasture. New Zealand Journal of Agricultural Research 36: 451-457.

Shackell, G., J. Wylie and R. Kelly. 1993b. Effects of prolonged exposure of ewes to oestrogenic pasture 2. Occurrence of abnormalities of the external genitalia and altered mating performance. New Zealand Journal of Agricultural Research 36: 459-464.

Sharma, S. 2017. Prebreeding using wild species for genetic enhancement of grain legumes at ICRISAT. Crop Science 57: 1132-1144. doi:10.2135/cropsci2017.01.0033.

Sharma, S., H.D. Upadhyaya, R.K. Varshney and C. Gowda. 2013. Pre-breeding for diversification of primary gene pool and genetic enhancement of grain legumes. Frontiers in Plant Science 4: 309. doi:10.3389/fpls.2013.00309.

Shaw, P.D., M. Graham, J. Kennedy, I. Milne and D.F. Marshall. 2014. Helium: visualization of large scale plant pedigrees. BMC Bioinformatics 15: 259. doi:10.1186/1471-2105-15-259.

Shaw, P.D., J. Kennedy, M. Graham, I. Milne and D.F. Marshall. 2016. Visualizing genetic transmission patterns in plant pedigrees. Edinburgh Napier University.

Sheaffer, C.C. and P. Seguin. 2009. Kura clover response to drought. Forage and Grazinglands 7: 0-0. doi:10.1094/FG-2009-1231-01-RS.

Sheath, G., M. Macfarlane and G. Crouchley. 1990. Evaluation of clovers in dry hill country 7. Subterranean and white clovers at Porangahau, Hawkes Bay, New Zealand. New Zealand Journal of Agricultural Research 33: 565-568. doi:10.1080/00288233.1990.10428458.

Shinozaki, K. and K. Yamaguchi-Shinozaki. 2007. Gene networks involved in drought stress response and tolerance. Journal of Experimental Botany 58: 221-227. doi:10.1093/jxb/erl164.

Shugart, Y.Y., Y. Zhu, W. Guo and M. Xiong. 2012. Weighted pedigree-based statistics for testing the association of rare variants. BMC Genomics 13: 667. doi:10.1186/1471-2164-13-667.

Sinclair, T. and M. Ludlow. 1986. Influence of soil water supply on the plant water balance of four tropical grain legumes. Functional Plant Biology 13: 329-341. doi:10.1071/PP9860329.

Sinclair, T.R., C.D. Messina, A. Beatty and M. Samples. 2010. Assessment across the United States of the benefits of altered soybean drought traits. Agronomy Journal 102: 475-482. doi:10.2134/agronj2009.0195.

Sinclair, T.R., M.A. Zwieniecki and N.M. Holbrook. 2008. Low leaf hydraulic conductance associated with drought tolerance in soybean. Physiologia Plantarum 132: 446-451. doi:10.1111/j.1399-3054.2007.01028.x.

Singh, N., S. Wu, W.J. Raupp, S. Sehgal, S. Arora, V. Tiwari, et al. 2019. Efficient curation of genebanks using next generation sequencing reveals substantial duplication of germplasm accessions. Scientific Reports 9: 650. doi:10.1038/s41598-018-37269-0.

Singh, S., P. Vikram, D. Sehgal, J. Burgueño, A. Sharma, S.K. Singh, et al. 2018. Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. Scientific Reports 8: 12527. doi:10.1038/s41598-018-30667-4.
Slatkin, M. 2008. Linkage disequilibrium - understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics 9: 477-485. doi:10.1038/nrg2361.
Slavov, G.T., R. Nipper, P. Robson, K. Farrar, G.G. Allison, M. Bosch, et al. 2014. Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass Miscanthus sinensis. New Phytologist 201: 1227-1239. doi:10.1111/nph.12621.
Sloane, R.J., R.P. Patterson and T.E. Carter Jr. 1990. Field drought tolerance of a soybean plant introduction. Crop Science 30: 118-123. doi:10.2135/cropsci1990.0011183X003000010027x.
Smith, J., S. Kresovich, M. Hopkins, S. Mitchell, R. Dean, W. Woodman, et al. 2000. Genetic diversity among elite sorghum inbred lines assessed with simple sequence repeats. Crop Science 40: 226-232. doi:10.2135/cropsci2000.401226x.
Smith, R., N. Taylor and S. Bowley. 1985. Red clover. Clover science and technology 25: 457-470.
Smýkal, P., C.J. Coyne, M.J. Ambrose, N. Maxted, H. Schaefer, M.W. Blair, et al. 2015. Legume crops phylogeny and genetic diversity for science and breeding. Critical Reviews in Plant Sciences 34: 43-104.
Sneller, C.H. 1994. Pedigree analysis of elite soybean lines. Crop Science 34: 1515-1522. doi:10.2135/cropsci1994.0011183X003400060019x.
Soleimani, V., B. Baum and D. Johnson. 2002. AFLP and pedigree-based genetic diversity estimates in modern cultivars of durum wheat [*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.]. Theoretical and Applied Genetics 104: 350-357. doi:10.1007/s001220100714.
Sonah, H., L. O'Donoughue, E. Cober, I. Rajcan and F. Belzile. 2015. Identification of loci governing eight agronomic traits using a GBS-GWAS approach and validation by QTL mapping in soya bean. Plant Biotechnology 13: 211-221. doi:10.1111/pbi.12249.
Song, M., W. Hao and J.D. Storey. 2015. Testing for genetic associations in arbitrarily structured populations. Nature Genetics 47: 550. doi:10.1038/ng.3244.
Souza, E. and M. Sorrells. 1989. Pedigree analysis of North American oat cultivars released from 1951 to 1985. Crop Science 29: 595-601. doi:10.2135/cropsci1989.0011183X002900030008x.
Speer, G. and D. Allinson. 1985. Kura clover (*Trifolium ambiguum*): Legume for forage and soil conservation. Economic Botany 39: 165-176.
Steketee, C.J., W.T. Schapaugh, T.E. Carter and Z. Li. 2020. Genome-wide association analyses reveal genomic regions controlling canopy wilting in soybean. G3: Genes, Genomes, Genetics 10: 1413-1425. doi:10.1534/g3.119.401016.
Suckling, F., M. Forde and W. Williams. 1983. Naturalised subterranean clover in New Zealand. New Zealand Journal of Agricultural Research 26: 35-43.
Sud, S., N.S. Bains and G.S. Nanda. 2005. Genetic relationships among wheat genotypes, as revealed by microsatellite markers and pedigree analysis. Journal of Applied Genetics 46: 375-379.
Sul, J.H., L.S. Martin and E. Eskin. 2018. Population structure in genetic studies: Confounding factors and mixed models. PLoS Genetics 14: e1007309. doi:10.1371/journal.pgen.1007309.
Sullivan, M.L. and R.D. Hatfield. 2006. Polyphenol oxidase and o-diphenols inhibit postharvest proteolysis in red clover and alfalfa. Crop Science 46: 662-670. doi:10.2135/cropsci2005.06-0132.
Sutherland, O.R., G.B. Russell, D.R. Biggs and G.A. Lane. 1980. Insect feeding deterrent activity of phytoalexin isoflavonoids. Biochemical Systematics and Ecology 8: 73-75. doi:10.1016/0305-1978(80)90029-0.
Sved, J. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theoretical Population Biology 2: 125-141. doi:10.1016/0040-5809(71)90011-6.
Takamizo, T., G. Spangenberg, K.-i. Suginobu and I. Potrykus. 1991. Intergeneric somatic hybridization in Gramineae: somatic hybrid plants between tall fescue (*Festuca arundinacea* Schreb.) and Italian ryegrass (*Lolium multiflorum* Lam.). Molecular and General Genetics 231: 1-6. doi:10.1007/BF00293814.
Tam, V., N. Patel, M. Turcotte, Y. Bossé, G. Paré and D. Meyre. 2019. Benefits and limitations of genome-wide association studies. Nature Reviews Genetics 20: 467-484. doi:10.1038/s41576-019-0127-1.

Tanksley, S.D. and S.R. McCouch. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277: 1063-1066.
Tashiro, R.M., Y. Han, M.J. Monteros, J.H. Bouton and W.A. Parrott. 2010. Leaf trait coloration in white clover and molecular mapping of the red midrib and leaflet number traits. Crop Science 50: 1260-1268. doi:10.2135/cropsci2009.08.0457.
Taylor, N., S. Ghabrial, S. Diachun and P. Cornelius. 1986. Inheritance and backcross breeding of the hypersensitive reaction to bean yellow mosaic virus in red clover. Crop Science 26: 68-74. doi:10.2135/cropsci1986.0011183X002600010016x.
Taylor, N. and E. Wiseman. 1985. Methodology and breeding of tetraploid red clover. Proceedings of the International Grasslands Congress 15: 244-245.
Taylor, N.L. 1982. Stability of S alleles in a doublecross hybrid of red clover Crop Science 22: 1222-1225.
Taylor, N.L. 1985. Clover science and technologyAmerican Society of Agronomy: Crop Science Society of America, Madison, Wisconsin, USA.
Taylor, N.L. 2008. A century of clover breeding developments in the United States. Crop Science 48: 1-13. doi:10.2135/cropsci2007.08.0446.
Taylor, N.L., P.L. Cornelius and R.E. Sigafus. 1984. Recurrent selection for forage and seed yield in zigzag clover. Canadian Journal of Plant Science 64: 119-130. doi:10.4141/cjps84-015.
Taylor, N.L., K. Johnston, M.K. Anderson and J.C. Williams. 1970. Inbreeding and heterosis in red clover. Crop Science 10: 522-525. doi:10.2135/cropsci1970.0011183X001000050021x.
Taylor, N.L. and K.H. Quesenberry. 1996a. Morphology and Physiology. Red Clover Science. Springer Netherlands. p. 44-56.
Taylor, N.L. and K.H. Quesenberry. 1996b. Tetraploid Red Clover. Red Clover Science. Springer Netherlands. p. 161-169.
Taylor, N.L. and R.R. Smith. 1980. Red clover breeding and genetics. Advances in Agronomy 31: 125-154. doi:10.1016/S0065-2113(08)60138-8.
Taylor, N.L. and R.R. Smith. 1997. Kura Clover (*Trifolium ambiguum* M.B.) Breeding, Culture, and Utilization. In: D. L. Sparks, editor Advances in Agronomy. Academic Press. p. 153-178.
Tenaillon, M.I., M.C. Sawkins, A.D. Long, R.L. Gaut, J.F. Doebley and B.S. Gaut. 2001. Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. mays L.). Proceedings of the National Academy of Sciences 98: 9161-9166. doi:10.1073/pnas.151244298.
Tessmann, E.W. and D.A. Van Sanford. 2018. GWAS for Fusarium head blight related traits in winter wheat (*Triticum aestivum* L.) in an artificially warmed treatment. Agronomy 8: 68. doi:10.3390/agronomy8050068.
Thomas, H. 1984. Effects of drought on growth and competitive ability of perennial ryegrass and white clover. Journal of Applied Ecology: 591-602. doi:10.2307/2403431.
Thompson, C.S. 1987. The climate and weather of the Hawke's Bay. Ministry of Transport, New Zealand Meteorological Service.
Townsend, C. 1964. Correlation among characters and general lack of persistence in diverse populations of alsike clover, *Trifolium hybridum* L. Crop Science 4: 575-577.
Townsend, C. and E. Remmenga. 1968. Inbreeding in Tetraploid Alsike Clover, *Trifolium hybridum* L. Crop Science 8: 213-217.
Tucak, M., S. Popovic, T. Cupic, V. Spanic and V. Meglic. 2013. Variation in yield, forage quality and morphological traits of red clover (*Trifolium pratense* L.) breeding populations and cultivars. Zemdirbyste 100: 63-70. doi:10.13080/z-a.2013.100.009.
Turkington, R. and J.L. Harper. 1979. The growth, distribution and neighbour relationships of *Trifolium repens* in a permanent pasture: I. Ordination, pattern and contact. The Journal of Ecology: 201-218. doi:10.2307/2259345.
Turley, P., R.K. Walters, O. Maghzian, A. Okbay, J.J. Lee, M.A. Fontana, et al. 2018. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nature Genetics 50: 229-237. doi:10.1038/s41588-017-0009-4.
Ulloa, O., F. Ortega and H. Campos. 2003. Analysis of genetic diversity in red clover (*Trifolium pratense* L.) breeding populations as revealed by RAPD genetic markers. Genome 46: 529-535. doi:10.1139/g03-030.

Umezawa, T., M. Fujita, Y. Fujita, K. Yamaguchi-Shinozaki and K. Shinozaki. 2006. Engineering drought tolerance in plants: discovering and tailoring genes to unlock the future. Current Opinion in Biotechnology 17: 113-122. doi:10.1016/j.copbio.2006.02.002.
Valera, M., A. Molina, J.P. Gutiérrez, J. Gómez and F. Goyache. 2005. Pedigree analysis in the Andalusian horse: population structure, genetic variability and influence of the Carthusian strain. Livestock Production Science 95: 57-66. doi:10.1016/j.livprodsci.2004.12.004.
van Berloo, R. and R.C.B. Hutten. 2005. Peditree: pedigree database analysis and visualization for breeding and science. Journal of Heredity 96: 465-468. doi:10.1093/jhered/esi059.
van den Bosch, J., I.K. Black, G.R. Cousins and D.R. Woodfield. 1993b. Enhanced drought tolerance in white clover. Proceedings of the New Zealand Grassland Association 55: 97-101.
van den Bosch, J., J. Lancashire, B. Cooper, T. Lyons and W. Williams. 1986. G18 white clover - a new cultivar for lowland pastures. Proceedings of the New Zealand Grassland Association 47: 173-177.
van Dijk, G. 1979. Wild species for the breeding of grasses. Broadening Genetic Base Crops. CABI. p. 211-216.
Van Ranst, G., M.R. Lee and V. Fievez. 2011. Red clover polyphenol oxidase and lipid metabolism. Animal 5: 512-521. doi:10.1017/S1751731110002028.
Van de Berg, J.L., M.J. Aivaliotis, L.E. Williams and C.R. Abee. 1990. Biochemical genetic markers of squirrel monkeys and their use for pedigree validation. Biochemical Genetics 28: 41-56. doi:10.1007/BF00554820.
Vanwijk, A.J.P. and D. Reheul. 1991. Achievements in fodder crops breeding in maritime Europe. In: A. E. A. Dennijis, editor Fodder Crops Breedings: Achievements, Novel Strategies and Biotechnology. The Netherlands. p. 13-18.
Varin, S., B. Leveel, S. Lemauviel-Lavenant and J.-B. Cliquet. 2009. Does the white clover response to sulphur availability correspond to phenotypic or ontogenetic plasticity? Acta Oecologica 35: 452-457. doi:10.1016/j.actao.2009.01.002.
Vaseva, I., Y. Akiscan, K. Demirevska, I. Anders and U. Feller. 2011. Drought stress tolerance of red and white clover–comparative analysis of some chaperonins and dehydrins. Scientia Horticulturae 130: 653-659. doi:10.1016/j.scienta.2011.08.021.
Verity, R. and R.A. Nichols. 2016. Estimating the number of subpopulations (K) in structured populations. Genetics 203: 1827-1839. doi:10.1534/genetics.115.180992.
Vineis, P. and N. Pearce. 2010. Missing heritability in genome-wide association study research. Nature Reviews Genetics 11: 589. doi:10.1038/nrg2809-c2.
Vinson, A., K. Prongay and B. Ferguson. 2013. The value of extended pedigrees for next-generation analysis of complex disease in the rhesus macaque. The Institute for Laboratory Animal Research Journal 54: 91-105. doi:10.1093/ilar/ilt041.
Virgona, J. and B. Dear. 1996. Comparative performance of Caucasian clover (*Trifolium ambiguum* cv. Monaro) after 11 years under low-input conditions in south-eastern Australia. New Zealand Journal of Agricultural Research 39: 245-253. doi:10.1080/00288233.1996.9513183.
Visscher, P.M., M.A. Brown, M.I. McCarthy and J. Yang. 2012. Five years of GWAS discovery. The American Journal of Human Genetics 90: 7-24. doi:10.1016/j.ajhg.2011.11.029.
Visscher, P.M., N.R. Wray, Q. Zhang, P. Sklar, M.I. McCarthy, M.A. Brown, et al. 2017. 10 years of GWAS discovery: biology, function, and translation. The American Journal of Human Genetics 101: 5-22. doi:10.1016/j.ajhg.2017.06.005.
Vleugels, T. 2013. Breeding for resistance to clover rot (*Sclerotinia* spp.) in red clover breeding (*Trifolium pratense*). Ghent University.
Volante, A., F. Desiderio, A. Tondelli, R. Perrini, G. Orasen, C. Biselli, et al. 2017. Genome-wide analysis of japonica rice performance under limited water and permanent flooding conditions. Frontiers in Plant Science 8: 1862. doi:10.3389/fpls.2017.01862.
Voorrips, R.E., M.C. Bink and W.E. van de Weg. 2012. Pedimap: software for the visualization of genetic and phenotypic data in pedigrees. Journal of Heredity. doi:10.1093/jhered/ess060.
VSN-International. 2019. Genstat for Windows 20th Edition. VSN International Ltd, Hemel Hempstead, UK.
Walton, P.D. 1971. The origin and development of world forage crops. Economic Botany 25: 263-266. doi:10.1007/bf02860763.
Ware, W. 1925. Experiments and observations on forms and strains of *Trifolium repens*. The Journal of Agricultural Science 15: 47-67.

Weisz, R., J. Kaminski and Z. Smilowitz. 1994. Water deficit effects on potato leaf growth and transpiration: utilizing fraction extractable soil water for comparison with other crops. American Potato Journal 71: 829-840. doi:10.1007/BF02849378.
Welham, C.V., R. Turkington and C. Sayre. 2002. Morphological plasticity of white clover (*Trifolium repens* L.) in response to spatial and temporal resource heterogeneity. Oecologia 130: 231-238. doi:10.1007/s004420100791.
Wery, J. 2005. Differential effects of soil water deficit on the basic plant functions and their significance to analyse crop responses to water deficit in indeterminate plants. Australian Journal of Agricultural Research 56: 1201-1209. doi:10.1071/AR05066.
White, J. and J. Hodgson. 1999. New Zealand Pasture and Crop Science.
Widdup, K. and B. Barrett. 2011. Achieving persistence and productivity in white clover. Pasture Persistence Symposium. Grassland Research and Practice Series: 173-180.
Widdup, K., J. Ford, G. Cousins, D. Woodfield, J. Caradus and B. Barrett. 2015. A comparison of New Zealand and overseas white clover cultivars under grazing in New Zealand. Journal of New Zealand Grasslands 77: 51-56.
Widdup, K., J. Garcia, J. Amadeo, R. Guillen and D. Real. 2006. White clover cultivars developed for temperate regions of South America. Proceedings of the 13th Australasian Plant Breeding Conference: 18-21.
Widdup, K., M. Hickey, D. Stevens and D. Ryan. 1989. A white clover bred for southern regions. Proceedings of the New Zealand Grassland Association 50: 207-212.
Widdup, K. and C. Pennell. 2000a. Suitability of new subterranean clovers in the Canterbury region. Proceedings of the New Zealand Grassland Association 62: 161-166.
Widdup, K. and D. Ryan. 1992. Forage potential of wild populations of perennial ryegrass collected from southern New Zealand farms. Proceedings of the New Zealand Grassland Association 54: 161-165.
Widdup, K.H., T.L. Knight and L.M. Hunt. 1996. Genetic variation for seed yield in Caucasian clover. Proceedings of the New Zealand Grassland Association: 189-194.
Widdup, K.H. and D.L. Ryan. 1994. Development of G50 alsike clover for the South Island high country. Proceedings of the New Zealand Grassland Association 56: 107-111.
Wiersma, D.W. 2001. Are hybrids the new yield force in alfalfa? A Summary of alfalfa hybrid performance in University Variety Trials: 1-4.
Wiggans, G.R., P.M. VanRaden and J. Zuurbier. 1995. Calculation and use of inbreeding coefficients for genetic evaluation of United States dairy cattle. Journal of Dairy Science 78: 1584-1590. doi:10.3168/jds.S0022-0302(95)76782-0.
Wilkins, P. and M. Humphreys. 2003. Progress in breeding perennial forage grasses for temperate agriculture. The Journal of Agricultural Science 140: 129-150. doi:10.1017/S0021859603003058.
Williams, R. and W. Williams. 1947. Genetics of red clover (*Trifolium pratense* L.) compatibility. Journal of Genetics 48: 51-68. doi:10.1007/BF02984421.
Williams, T., M. Abberton, W. Thornley, D. Evans and I. Rhodes. 1998. Evaluation of seed production potential in white clover (*Trifolium repens* L.) varietal improvement programmes. Grass and Forage Science 53: 197-207. doi:10.1046/j.1365-2494.1998.5320197.x.
Williams, W. 1951. Genetics of incompatibility in alsike clover (*Trifolium hybridum*). Heredity 5: 51. doi:10.1038/hdy.1951.3.
Williams, W. 2000. Germplasm centres and issues of seed conservation. Seed Symposium 12: 103-107.
Williams, W. 2010. The key roles of seed banks in plant biodiversity management in New Zealand. Seed symposium: Seeds for Futures: 5-11.
Williams, W. 2014. *Trifolium* interspecific hybridisation: widening the white clover gene pool. Crop and Pasture Science 65: 1091-1106. doi:10.1071/CP13294.
Williams, W. 1987. Genetics and breeding. In: M. J. Baker and W. M. Williams, editors, White clover. CAB Int, Wallingford, UK. p. 343-420.
Williams, W. and J. Caradus. 1979. Performance of white clover lines on New Zealand hill country. Proceedings of the New Zealand Grassland Association 40: 162-169.
Williams, W., H. Easton and C. Jones. 2007. Future options and targets for pasture plant breeding in New Zealand. New Zealand Journal of Agricultural Research 50: 223-248. doi:10.1080/00288230709510292.

Williams, W., I. Verry and N. Ellison. 2006. A phylogenetic approach to germplasm use in clover breeding. Proceedings of the 13th Australasian Plant Breeding Conference: 966-971.
Williams, W., I. Verry, S. Hussain, H. Ansari, K. Widdup, N. Ellison, et al. 2013. Widening the adaptation of white clover by incorporation of valuable new traits from wild clover species. Proceedings of the 22nd International Grasslands Congress: 15-19.
Williams, W.M., N.W. Ellison, H.A. Ansari, I.M. Verry and S.W. Hussain. 2012. Experimental evidence for the ancestry of allotetraploid *Trifolium repens* and creation of synthetic forms with value for plant breeding. BMC Plant Biology 12: 55. doi:10.1186/1471-2229-12-55.
Williams, W.M. and S.W. Hussain. 2008. Development of a breeding strategy for interspecific hybrids between Caucasian clover and white clover. New Zealand Journal of Agricultural Research 51: 115-126.
Williams, W.M., I.M. Verry, H.A. Ansari, S.W. Hussain, I. Ullah and N.W. Ellison. 2019. A Eurasia-wide polyploid species complex involving 6x *Trifolium ambiguum*, 2x *T. occidentale* and 4x *T. repens* produces interspecific hybrids with significance for clover breeding. BMC Plant Biology 19: 438. doi:10.1186/s12870-019-2030-5.
Williams, W.M., I.M. Verry, H.A. Ansari, S.W. Hussain, I. Ullah, M.L. Williamson, et al. 2011. Eco-geographically divergent diploids, Caucasian clover (*Trifolium ambiguum*) and western clover (*T. occidentale*), retain most requirements for hybridization. Annals of Botany 108: 1269-1277. doi:10.1093/aob/mcr226.
Wilsie, C. 1958. Effect of Inbreeding on Fertility and Vigor of Alfalfa. Agronomy Journal 50: 182-185. doi:10.2134/agronj1958.00021962005000040004x.
Wisser, R.J., S.C. Murray, J.M. Kolkman, H. Ceballos and R.J. Nelson. 2008. Selection mapping of loci for quantitative disease resistance in a diverse maize population. Genetics 180: 583-599. doi:10.1534/genetics.108.090118.
Woodfield, D. and E. Brummer. 2001a. Integrating molecular techniques to maximise the genetic potential of forage legumes. In: G. Spangenberg, editor Molecular Breeding of Forage Crops. Developments in Plant Breeding. Springer, Dordrecht. p. 51-65.
Woodfield, D., P. Clifford, I. Baird, G. Cousins, J. Miller, K. Widdup, et al. 2003. Grasslands Tribute: a multi-purpose white clover for Australasia. Proceedings of the New Zealand Grassland Association 65: 157-162.
Woodfield, D., P. Clifford, G. Cousins, J. Ford, I. Baird, J. Miller, et al. 2001b. Grasslands Kopu II and Crusader: new generation white clovers. Proceedings of the Conference - New Zealand Grassland Association: 103-108.
Woodfield, D., J. Ford, M. Jahufer and D. Johnston. 2006. New generation white clovers for United Kingdom farming systems. Proceedings of the 13th Australasian Plant Breeding Conference: 66-70.
Woodfield, D.R. 1999. Genetic improvements in New Zealand forage cultivars. Proceedings of the Conference - New Zealand Grassland Association 61: 3-8.
Woodfield, D.R. and J.R. Caradus. 1994. Genetic improvement in white clover representing six decades of plant breeding. Crop Science 34: 1205-1213. doi:10.2135/cropsci1994.0011183X003400050011x.
Woodfield, D.R. and J.R. Caradus. 1996. Factors affecting white clover persistence in New Zealand pastures. New Zealand Grassland Association: 229-236.
Wratt, G.S. and H.C. Smith. 2015. Plant Breeding in New ZealandElsevier Science.
Wricke, G. and E. Weber. 1986. Quantitative Genetics and Selection in Plant Breeding. Walter de Gruyter.
Wright, S. 1984a. Evolution and the genetics of populations: Experimental results and evolutionary deductions. University of Chicago Press 3.
Wright, S. 1984b. Experimental Results and Evolutionary Deductions. University of Chicago Press.
Wu, S., J. Yang and R. Wu. 2010. Mapping quantitative trait loci in a non-equilibrium population. Statistical Applications in Genetics and Molecular Biology 9. doi:10.2202/1544-6115.1578.
Xing, Y., U. Frei, B. Schejbel, T. Asp and T. Lübberstedt. 2007. Nucleotide diversity and linkage disequilibrium in 11 expressed resistance candidate genes in *Lolium perenne*. BMC Plant Biology 7: 43. doi:10.1186/1471-2229-7-43.
Xiong, L., R.-G. Wang, G. Mao and J.M. Koczan. 2006. Identification of drought tolerance determinants by genetic analysis of root response to drought stress and abscisic acid. Plant Physiology 142: 1065-1074. doi:10.1104/pp.106.084632.

Yamada, M. and T. Hasegawa. 1990. New forage crop varieties registered by the Ministry of Agriculture: Forestry and Fisheries in 1989 and 1990. Japanese Journal of Breeding 40: 549-554.

Yamada, T., H. Fukuoka and T. Wakamatsu. 1989. Recurrent selection programs for white clover (*Trifolium repens* L.) using self-compatible plants. I. Selection of self-compatible plants and inheritance of a self-compatibility factor. Euphytica 44: 167-172.

Ye, H., L. Song, W.T. Schapaugh, M.L. Ali, T.R. Sinclair, M.K. Riar, et al. 2020. The importance of slow canopy wilting in drought tolerance in soybean. Journal of Experimental Botany 71: 642-652. doi:10.1093/jxb/erz150.

Yendrek, C.R., T. Tomaz, C.M. Montes, Y. Cao, A.M. Morse, P.J. Brown, et al. 2017. High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. Plant Physiology 173: 614-626. doi:10.1104/pp.16.01447.

You, M., M. Barbett and P. Nichols. 2005a. New sources of resistance identified in *Trifolium subterraneum* breeding lines and cultivars to root rot caused by Fusarium avenaceum and Pythium irregulare and their relationship to seedling survival. Australasian Plant Pathology 34: 237-244. doi:10.1071/AP04092.

You, M., M. Barbetti and P. Nichols. 2005b. New *Trifolium subterraneum* genotypes identified with resistance to race 2 of *Kabatiella caulivora* and cross-resistance to fungal root rot pathogens. Australian Journal of Agricultural Research 56: 1111-1114. doi:10.1071/AR05103.

Yu, J. and E.S. Buckler. 2006. Genetic Association Mapping and Genome Organization of Maize. Current Opinion in Biotechnology 17: 155-160. doi:10.1016/j.copbio.2006.02.003.

Yu, L.-X. 2017. Identification of single-nucleotide polymorphic loci associated with biomass yield under water deficit in alfalfa (*Medicago sativa* L.) using genome-wide sequencing and association mapping. Frontiers in Plant Science 8: 1152. doi:10.3389/fpls.2017.01152.

Yu, X., X. Li, T. Guo, C. Zhu, Y. Wu, S.E. Mitchell, et al. 2016. Genomic prediction contributing to a promising global strategy to turbocharge gene banks. Nature Plants 2: 16150. doi:10.1038/nplants.2016.150.

Zamir, D. 2001. Improving plant breeding with exotic genetic libraries. Nature Reviews Genetics 2: 983. doi:10.1038/35103590.

Zeven, A. 1991. Four hundred years of cultivation of Dutch white clover landraces. Euphytica 54: 93-99. doi:10.1007/BF00145635.

Zhang, H., R. Bai, F. Wu, W. Guo, Z. Yan, Q. Yan, et al. 2019. Genetic diversity, phylogenetic structure and development of core collections in Melilotus accessions from a Chinese gene bank. Scientific Reports 9: 1-9. doi:10.1038/s41598-019-49355-y.

Zhang, Y., C. Motes, M.K. Sledge, J.H. Bouton, Y. Han and M.J. Monteros. 2010. Identification of QTLs associated with morphological and agronomic traits in white clover (*Trifolium repens* L.). Sustainable Use of Genetic Diversity in Forage and Turf Breeding. Springer. p. 489-492.

Zhang, Y., M.K. Sledge and J.H. Bouton. 2007. Genome mapping of white clover (*Trifolium repens* L.) and comparative analysis within the *Trifolieae* using cross-species SSR markers. Theoretical and Applied Genetics 114: 1367-1378. doi:10.1007/BF00017991.

Zhao, K., C.-W. Tung, G.C. Eizenga, M.H. Wright, M.L. Ali, A.H. Price, et al. 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. Nature Communications 2: 467. doi:10.1038/ncomms1467.

Zheng, J., H. Liu, Y. Wang, L. Wang, X. Chang, R. Jing, et al. 2014. TEF-7A, a transcript elongation factor gene, influences yield-related traits in bread wheat (*Triticum aestivum* L.). Journal of Experimental Botany 65: 5351-5365. doi:10.1093/jxb/eru306.

Zhou, H., J. Zhou, E.M. Sobel and K. Lange. 2014. Fast genome-wide pedigree quantitative trait loci analysis using MENDEL. BMC Proceedings 8: S93. doi:10.1186/1753-6561-8-S1-S93.

Zohary, M. and D. Heller. 1984. The genus *Trifolium*. Israel Academy of Sciences and Humanities.