Persistency and genomic tools in forage grasses

Niamh Curran
March 2023

Thesis submitted in fulfilment of the requirements for the Doctor of Philosophy at Trinity College Dublin

Supervisors:

Prof. Trevor R. Hodkinson
Department of Botany, School of Natural Sciences, Trinity College Dublin

Dr. Dan Milbourne
Senior Research Officer, Teagasc Crops Research Centre, Oak Park, Carlow

Dr. Susanne Barth
Principal Research Officer, Teagasc Crops Research Centre, Oak Park, Carlow

Co-supervisor:
Dr. Stephen L. Byrne
Research Officer, Teagasc Crops Research Centre, Oak Park, Carlow

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

_____

Niamh Curran
June 2021

Acknowledgements

This research was funded through the Walsh Scholarship programme. I would like to acknowledge the guidance and mentorship of my supervisors, Dr Dan Milbourne, Dr Susanne Barth, Dr Stephen Byrne and Dr Trevor Hodkinson.

I would like to acknowledge the staff of Teagasc, Oakpark. In particular, I would like to thank Dr Aude Perdereau, Dr Maria Leyva, Dr Fergus Meade and Dr Helena Mealy for their help and advice in the lab and in data analysis. I would like to thank Dr Michael O'Donovan and Dr Tomas Tubritt for all of their work maintaining trials and facilitating sample collection. I would like to acknowledge the support and friendship provided by my peers, Leona, Katie, Tomas, Tara and Niall.

I would like to thank my family for all of their help and kindness throughout my academic career, namely Rosarii, Ruairi and Jack. Finally, I would like to thank Ronan for his encouragement and patience during the completion of this thesis.

Summary

Grasslands are integral to sustainable ruminant production. Perennial ryegrass (PRG; *Lolium perenne*) is the most widely cultivated forage grass species in Ireland. Despite its widespread use, there have been meagre gains in the improvement of agronomically important traits. This is due to the only recent progress in trait enhancement, in comparison to other agriculturally important species, as well as the approaches undertaken in the breeding of improved cultivars. Persistency is a trait which is of economic importance but has lacked improvement. Persistency in the context of forage breeding, refers to the expected performance of a sward over time. Novel breeding techniques such as hybridisation and genomic approaches have the potential to advance the understanding of persistency as well contribute to the improvement of the trait. Genomic approaches such as reduced representation assays can effectively accelerate the characterisation of PRG diversity as well understand how populations change in the sward over time. Swards of mixed cultivars are sown due to the perception of synergism between cultivars of complementary characteristics. Chapter 1 discusses the different aspects of persistency as an economically important trait, its improvement, the current methods of quantification and the capacity to utilise molecular marker to measure and improve persistency. The genomic resources and approaches currently utilised in the improvement of forages, current genotyping methods available and their advantages and limitations are also discussed.

The *Lolium* and *Festuca* genera are attractive candidates for hybridisation due to their complementary characteristics in the context of forage production and environmental challenges, with *Lolium* exhibiting greater digestibility and *Festuca* possessing superior resilience to abiotic stress. In Chapter 2, we aimed to establish the extent of the representation of these characteristics in different *Festulolium* hybrids, a five year field experiment was undertaken in a mild Atlantic climate. Overall, *Festulolium* hybrids performed at a very similar level as *Lolium* species in terms of persistency, yield and ensilibility.

In Chapter 3 the design of a low-cost high throughput genotyping platform, GenoGrass, was undertaken to investigate the utility of reduced representation

assays in population delineation in PRG and detecting change in allele frequencies in PRG mixtures. This was based on genotyping-in-thousands by sequencing (GT-seq) and relies on a two fold polymerase chain reaction (PCR) to amplify target regions. A total of 192 sites (loci) were selected for primer design, 54 of which were identified from previous research which had high predictive ability for heading date and crown rust resistance, while the remaining sites were selected for their even distribution across the PRG genome. This assay was then deployed in a number of PRG populations with varying levels of complexity. Reduced representation assays did delineate PRG populations and performed comparably to approaches with a much greater number of SNPs. However, some aspects of the PRG genome, including repetitive regions, as well as the incomplete nature of PRG genomic resources likely led to technical challenges in the deployment of this assay. We envisaged a revised version of this assay being useful in PRG for identifying populations and the variants driving population delineation and subsequently providing markers for persistency.

In Chapter 4 an alternative reduced representation genotyping approach was implemented (SeqSNP). We investigated the changes in mixture components over time, reflective of mixtures sown on Irish farms. Competition and selective pressures present in the field have the potential to affect the survival of individual plants in the sward and therefore alter proportion of mixture components. Adjustments were made to the genotyping approach, following the results of Chapter 3. This was to mitigate the issue encountered previously. Some shifts in genetic profiles were observed from the original seed however these changes were sporadic and no clear trends or relationships were observed.

In summary, reduced representation assays efficiently and adequately captured genetic diversity in PRG pooled populations. These assays can be used for many applications in PRG breeding. Swards of mixed cultivars exhibit little proportional changes over a two year period. Though this study took place over a brief period in the life time of a sward, it gives a good indication of sward dynamics, post sowing and in the subsequent two years. Chapter 5 summarises how this thesis has added to the growing knowledge of persistency in forage grasses and developed and evaluated tools for its study - in addition to wider applications in plant breeding.

Future studies should incorporate the findings of this investigation, including the implications for genomic resources and measures of persistency.

Contents

Chapter 1

General introduction

1.1 Grass in Irish agriculture

The agrifoods sector represents the largest indigenous industry of the Irish economy and due to the large export market for dairy and beef, ruminant production represents a significant portion of Irish economic output. The Irish dairy sector has undergone recent expansion with abolition of EU-instated milk quotas in 2015 (Kelly *et al*., 2020). The increasing dairy herd is at odds with climate mitigation commitments due to the associated greenhouse gas emissions attributed to ruminant production (Kelly *et al*., 2020). This, coupled with unstable markets can negatively impact profitability and production in this industry. Grass is a cheap, high quality source of feed for ruminants and rain-fed grass production systems are the primary source of feed in Irish beef and dairy production (Finneran *et al*., 2010). The main forage grasses in Ireland include species within the genera *Lolium*, *Festuca*, *Dactylis*, *Phleum* and *Poa*. Despite the studied benefits of multispecies swards, single perennial ryegrass (PRG) species swards with marginal inclusion of white clover dominate grass seed sales (DAFM, 2021). Grazed grass comprises 70% of the diet of a bovine animal, with grass based silage constituting a further 23% and the remaining 7% comprised of concentrates (Shalloo *et al*., 2004). Although ruminant agriculture is a contributor to greenhouse gas (GHG) emissions, grasslands potentially offset such emission through carbon sequestration (O'Mara, 2012). Maximizing the proportion of grazed grass in the diet for dairy and beef cattle has been demonstrated to minimise production costs. This will be an integral part of fulfilling the demands of an expanding herd while mitigating the effects of an unstable market and climate change (O'Mara, 2012; Dillon *et al*., 2016).

1.1.1 Perennial ryegrass

PRG (*Lolium perenne*) is one of the most important forage grass species in temperate regions worldwide (Wilkins & Humphreys, 2003). PRG is widely cultivated due to its response to inputs and regrowth capacity under frequent defoliation (Tallowin *et al*., 1995). This is essential for intensive grazing systems used in Irish beef and dairy production, and PRG represents over 95% of the grass seed sold in Ireland (DAFM, 2021). As its name suggests, PRG is perennial in nature, and through the asexual reproduction of the tiller swards it can often remain in the ground for over a decade before reseeding is carried out. PRG is an obligate

outbreeder due to self-incompatibility (Manzanares *et al*., 2016), and populations are produced from random intercrossing of multiple parents which have undergone prior selection in swards and as single plants. Consequently, PRG cultivars represent heterogeneous populations with considerable within population variability (Barth *et al.,* 2017). PRG is naturally diploid while tetraploid varieties have been generated using colchicine (Myers, 1939). Improvement of PRG began recently in comparison to some other agriculturally important species and phenotypic recurrent breeding is traditionally used to improve agriculturally important traits in PRG (Wilkins, 1991). Labour-intensive phenotypic selection and lengthy completion time has led to modest gains in trait improvement. Productivity has seen the greatest improvement in genetic gain within the last 40 years (McDonagh *et al*., 2016). Genetic gain in this context refers to the improvement of agriculturally beneficial traits (Xu *et al*., 2017). As the end product in grassland cultivation is beef and milk, the nutritional content and palatability to ruminants are also key traits (McGrath *et al.,* 2014). Increases in nutritive values have also seen modest improvements over the same period. Persistency is associated with yield stability and despite being an economically important trait, has not seen the same level of improvement (McDonagh *et al*., 2016). The concept of persistency is further explained in Section 1.2 of this Chapter.

1.1.2 Fescues and hybrids

The *Festuca-Lolium* complex represents the closely related species of the *Lolium* and *Festuca* genera (Thomas *et al*., 2003) in the *Pooideae* subfamily (Hodkinson *et al.*, 2018). Species from this complex possess physiological similarities, however, agronomically, fescues offer distinct positive attributes in comparison to *Lolium* species. The root physiology of *Festuca* has been attributed to its ability to withstand abiotic stress (Huang & Gao, 2000). Tall fescue has been observed to increase root growth under drought conditions, increase water uptake in the lower soil and maintain root activity (Huang & Gao, 2000). *Festuca* have been observed to tolerate drought and cold well, in comparison to other grass species (Rognli *et al*., 2010). The root physiology of *Festuca* has also been observed to improve soil structure and porosity, alleviating negative effects of flooding (Macleod *et al*., 2013). This may improve the resilience of grasslands in the instance of more frequent extreme

weather events brought about by climate change. However, fescues are comparatively less palatable and tolerant to grazing than ryegrasses. *Festuca* and *Lolium* possess the natural ability to hybridise, with hybrids collectively known as *Festulolium*. Intergeneric hybridisation has the potential to combine beneficial traits of both species (Humphreys & Zwierzykowski, 2020). This approach has been implemented as a means of improving persistency and abiotic stress tolerance in *Lolium* species while maintaining positive grazing attributes.

1.2 Persistency

'Persistence' is a complex trait that refers to the expected performance of a sown sward over time (Tozer *et al*., 2014). The relationship with production makes persistency an economically important trait, recognised in cultivar evaluation, certification and the Teagasc Pasture Profit Index, an economic ranking of perennial ryegrass cultivars (Grogan & Gilliland, 2011; O'Donovan *et al*., 2017). Persistency has many contributing factors as cultivars must withstand stresses present under farm conditions including winter cold stress, drought stress, pest damage, intense defoliation and treading damage (Clark, 2011). Such selective pressures acting on heterogenous populations have the potential to alter sward productivity and composition. Vegetative reproduction is the main mode of propagation, and due to the defoliation regimes of managed PRG swards propagation through seed is rare. This results in little sexual recombination within the sward and natality vs mortality of tillers determines the survival of grass within the sward (Duchini *et al*., 2018).

As the rate of reseeding increases so does the herbage mass of the sward (Shalloo et al., 2011). In Ireland, only 1-5% of grassland is reseeded annually (Humphreys & Casey, 2002; Creighton *et al*., 2011), significantly less than of the optimum rate of 8-10% to maximise efficiency (Shalloo *et al*., 2011). Reseeding is a costly practice due to the expense of seed, preparation and sowing as well as the inhibition of effective land utilisation during sward establishment. Improved persistency increases long term productivity without the need to increase the frequency of reseeding, minimising costs (Shalloo *et al*., 2011).

The definition of persistence varies in the literature with some authors focusing on the survival and persistence of plants and the abundance of species in the sward while others interpreting the trait as the persistence of yield rather than the individual

plant (Camlin & Stewart 1978; Chapman *et al*., 2015). Different methods of measurement quantify the different aspects of persistency (Dodd *et al.*, 2018). Ground score (GS) is the visual estimation of the proportion of PRG present in the sward (Charles & Valentine, 1978). Cultivar trials use GS to determine the persistency of a cultivar in monoculture for recommended seed lists. It is a quick, efficient measure of the abundance of PRG tillers. Ground score change (GSΔ), evaluating the decline in percentage of ground covered by sown species, is another measure of grass persistency (Clouting & Hawkins, 1966). However, both methods are considered subjective "point-in-time" measurements which may not be a true representation of persistence during the lifetime of a cultivar under animal grazing. These measures mostly quantify the abundance element of persistency, and can only be used to measure the overall persistency of the sward and not the contributions of individual cultivars. Measures which consider the relationship of persistence to yield, more accurately quantify the trait in the context of agronomic performance (Dodd *et al*., 2018). In previous Teagasc research, O'Donovan *et al*. (2016) conducted a study examining the quantification of persistency in PRG varieties. The relationship between dry matter (DM) yield data and GSΔ was used to indicate the useful lifespan of PRG varieties. This method was found to have limited use in accurately determining cultivar persistence because of the difficulties associated with estimating persistency in the use of short-term cutting plots as the short-term performance has to be extrapolated to predict longer term persistency. Investigating the changes in populations over time by using a genomic approach can help determine if decline in persistency over time is due to a loss of individual plants and therefore a loss in genetic diversity, or rather a loss of tillering with genetic diversity remaining relatively stable.

In this thesis, the concept of persistency as an individual plant's ability to survive in the sward will be explored. In Chapter 2 the utility and limitations of GS as a measure of persistency will be investigated. Persistency as allele frequencies and the changes over time to better understand the populations changes that occur will be the subject of Chapter 4.

1.3 Mixtures

It is widespread practice to sow PRG swards in complex mixtures (either multispecies or conspecific). In all of these systems, selective pressure can unevenly affect population structure. This is exacerbated by the already heterogeneous nature of PRG populations. This is particularly the case when cultivars have differing levels of persistency. Synergism can be observed in mixtures as “overyielding” of PRG swards, whereby the sowing of mixed cultivars results in a higher yield than that of the highest yielding component (Nyfeler *et al*., 2009). Recently, complex mixtures were investigated by Tubritt *et al.* (2021) for synergism by comparing the aspects of productivity in mixtures with that of each of the mixture components. Limited and inconsistent synergism was observed in traits such as yield and utilisation. Furthermore, rather than synergism, a dilution effect was observed, where lower yielding cultivars reduced the overall yield of the mixture. As well as an unclear contribution, sward instability can affect the agronomic performance of the sward. Morphological similarities between cultivars result in difficulties when discerning the contribution of each component within mixture (Quaite & Camlin, 1986) and molecular methods are therefore required for this purpose. Verwimp *et al*. (2018) demonstrated how a molecular approach can be used to characterise and temporal dynamics of multispecies swards as well as single species mixed cultivar PRG swards. Verwimp *et al.* (2018) investigated the changes in yield and biodiversity in bispecific and monospecific plots of PRG and white clover with different growth habits. The approach of pooling DNA from PRG populations to track the changes in allele frequencies over time was validated. Some changes in the population composition of the PRG cultivar mixtures were detected over a 4 year field study, indicating the dynamic nature of sward composition over time.

1.4 Cultivar evaluation and certification

Regulatory tests exist to protect both breeders and consumers and to maintain genetic gain in forage grass species (Gilliland & Gensollen, 2010). Distinctness, uniformity and stability (DUS) tests have been put in place as a regulatory procedure within the European Union to ensure new cultivars are distinguishable from existing ones. This protects the intellectual property of seed breeders as well as guaranteeing the improvement of the cultivated species for farmers. Furthermore,

value of cultivation and use (VCU) tests exist to ensure that new cultivars offer an improvement on existing varieties (Grogan & Gilliland, 2011). In the case of DUS, guidelines are set out by The International Union for the Protection of New Varieties of Plants (UPOV), while international guidelines have not been collated for VCU tests. Numerous phenotypic traits are evaluated over several years to validate cultivars in DUS testing, at several sites. They require extensive phenotypic characterisation in comparison to existing cultivars. The inclusion of molecular markers potentially improve efficacy and enhance the regulatory process (Gilliland *et al*., 2019).

1.5 Genomics in forages

Genomics has allowed for an acceleration in the improvement of economically important traits in many crop species (Bevan *et al*., 2017). Though genetic gain has been comparatively low in comparison to other agronomically important species, the implementation of genomic approaches in PRG research and breeding has the potential to improve economically important traits (Byrne *et al*., 2017; Arojju *et al*., 2018; Arojju *et al*., 2020). Advances in genotyping and reduction in costs has improved the accessibility of genomic selection which has the potential to accelerate the increase of genetic gain in PRG (Hayes *et al*., 2013). Several marker types have been widely used (Jones *et al*., 2009) in PRG genomics including Random Amplified Polymorphic DNA (RAPDs), Simple Sequence Repeats (SSRs) and Chloroplast Simple Sequence Repeats (cpSSRs) (Bolaric *et al*., 2005; Diekmann *et al*., 2012; Barth *et al*., 2017). Jones *et al.* (2009), summarises the advantages and limitations of these different marker types. RAPDs are simple, inexpensive and no prior knowledge of the target genome is necessary, though reproducibility and reliability are a limitation. SSR markers are a PCR based marker system, they are effective in distinguishing closely related individuals and are abundant in plant genomes, they require significant developmental time. Single nucleotide polymorphisms (SNPs) are flexible codominant marker types which offer several technical advantages over alternative marker systems such as reproducibility and transferability between species. They have been used to enhanced genomic resources and to inform and improve approaches (Mammadov *et al*., 2012). Genotyping-by-sequencing (GBS) is a widely used approach due to its ability to simultaneously screen large numbers

of SNPs (100,000s) at a relatively low cost (Elshire *et al*., 2011). This approach has been useful for breeding applications as well as characterising pooled populations (Byrne *et al*., 2013; Verwimp *et al*., 2018). Two PRG reference genomes have been assembled, with one reference genome developed utilising synteny with other grass species, *Brachypodium, Oryza sativa ,* and *Sorghum bicolor* (Pfeifer *et al*., 2013; Byrne *et al*., 2015; Velmurugan *et al*., 2016). Both are unanchored and incomplete, however in draft they have been successfully utilised as a resource in many approaches for PRG improvement. Despite the progress in sequencing technologies and a reduction in marker generation costs, whole genome sequencing can still be cost prohibitive and unsuitable for many applications. Targeted resequencing approaches capturing a low to moderate quantity of sites (100s-1000s) while maintaining depth of coverage and reducing costs are required. Such approaches include NextSeq Series Amplicon Sequencing by Illumina, SeqSNP by LGC and GT-seq (Campbell *et al*., 2015). They require *a priori* knowledge of targeted site, however they also capture novel variation.

1.6 Aims and objectives

The overall aim of this thesis was to investigate the trait of persistency in swards reflective of that sown on Irish farms. We aimed to develop a high throughput genomic tool to effectively and cheaply detect changes in allele frequencies in PRG populations. We aimed to deploy this in detecting changes in complex PRG mixtures. This will provide evidence for the selection and development of PRG mixtures for Irish farmers. The specific objectives of the thesis were to:

- Evaluate persistency of several important grass species including *Lolium*, *Festuca* and hybrids over an extended period and investigate the potential of the improvement of persistency through hybridisation (Chapter 2)
- Develop a high-throughput low cost genotyping platform (GenoGrass) to detect subtle changes in allele frequencies in PRG populations, cultivars and mixed populations in a pilot study (Chapter 3)
- Based on the high throughput genotyping pilot study, develop an approach to characterise PRG mixtures and discern changes in allele frequencies in swards of mixed PRG cultivars over time (Chapter 4)
- Synthesise the results from each chapter to reach conclusions about persistency of forage grass cultivars in Ireland and the methods used to characterised persistency (Chapter 5; final discussion)

Chapter 2

Persistency, yield, and silage quality of *Festulolium* cultivars over a consecutive 5 year period, under a mild Atlantic climate

Author contribution

Dermot Grogan and Susanne Barth conceived and designed study. Dermot Grogan, E O'Riordan and M Hanley performed the sample collection. Jim Grant and I performed the data analysis. I led the authoring of the manuscript and interpretation of the data. Susanne Barth, Stephen L Byrne, Dan Milbourne and Trevor R Hodkinson contributed to interpretation of data and preparation of the final manuscript. All authors read and approved the final version.

Status

This has been published in a special issue of Biologia Plantarum

Citation

Curran, N., Grogan, D., Milbourne, D., Byrne, S.L., O'Riordan, E., Hanley, M., Grant, J., Hodkinson, T.R. and Barth, S., 2021. Persistency, yield, and silage quality of Festulolium cultivars over a consecutive five-year period under a mild Atlantic climate. Festulolium: From the Nature to Modern Breeding, p.89-108.

Edited by D Kopecký & A Kosmala

Abstract

Festulolium are hybrids between species within the *Lolium* and *Festuca* genera. They are attractive candidates for hybridisation due to their complementary characteristics in the context of forage production and environmental challenges, with *Lolium* exhibiting greater digestibility and *Festuca* possessing superior resilience to abiotic stress. To establish the extent of the representation of these characteristics in different *Festulolium* hybrids, a five year field experiment was undertaken in a mild Atlantic climate. Four alternative species combinations and six pure species controls were evaluated for dry matter yield and persistency from 2014 to 2018, while various nutritional parameters in first silage cut were measured from 2015 to 2017. *Festulolium* hybrids derived from *L. multiflorum* were found to group with both *Lolium* and *Festuca* pure species controls for mean ground score to evaluate persistency ($p<0.05$). Hybrid types of the *L. perenne* lineage were found to have ground score values more alike to their *Lolium* component, and distinct from the *F. pratensis* pure species control. In terms of dry matter yield, the majority of *L. multiflorum* hybrids observed formed a higher yielding group with the tetraploid *L. multiflorum* pure species control and distinct from diploid *F. pratensis* pure species control, while all other hybrid types and controls shared groups in common. For the nutrient content parameters including dry matter digestibility, water soluble carbohydrates and buffering capacity, *Lolium* controls exhibited more favourable values, while hybrids were largely found to display intermediate to low values. *Festuca* pure species controls consistently exhibited poorer values with the exception of crude protein content where *F. pratensis* displayed the highest values and the 2017 measurement of buffering capacity. In conclusion, the *Festulolium* hybrids shared characteristics with both parental types but they often showed greater likeness to the *Lolium* component.

Key words: *Festuca*, *Lolium*, ground cover, yield, nutritional value, persistency

2.1 Introduction

Forages such as *Lolium* and *Festuca* represent a low cost source of feed in ruminant agriculture in temperate regions worldwide (Wilkins & Humphreys, 2003). Areas in Europe which traditionally support forages are projected to experience increased instances and intensities of extreme weather such as drought and flooding due to climate change (Roudier *et al.,* 2016). Forages with wide adaptability and also the productivity potential to support intensive grazing systems are required to meet these challenges (Lin, 2011; O'Donovan *et al*., 2011). *Festulolium* are natural and artificial interspecific hybrids derived from the genera *Festuca* and *Lolium*, and species from each of these genera exhibit such complementary characteristics in the context of forage production, qualifying them as suitable candidates for hybridisation (Thomas & Humphreys, 1991). *Lolium* species are widespread in ruminant agriculture in temperate regions globally, with Italian ryegrass (*L. multiflorum*) and perennial ryegrass (*L. perenne*) cultivars being most widely sown (Wilkins and Humphreys, 2003). Perennial ryegrass cultivars are considered high performance forage varieties, withstanding extended periods of grazing and demonstrating favourable palatability and productivity in response to inputs. Italian ryegrass is a biennial forage crop with rapid establishment but comparatively lower persistency, suited for short term swards. Fescues are not noted for their agronomic performance; they exhibit poorer palatability and lower digestibility relative to ryegrasses. Fescues, however, display greater resistance to abiotic stress, such as drought and flooding, attributed to well-developed root systems. Meadow fescue (*F. pratensis*) or tall fescue (*F. arundinacea*) are the most widely sown *Festuca* species for forage, typically utilised species rich permanent pasture.

The level of integration of genomic material in hybrids is dependent on breeding strategy, through the generation of an amphiploid hybrid or by way of introgression. The former entails each parent donating at least one pair of chromosomes, though often resulting in *Lolium* chromatin dominating the proportion of genetic material (Kopecký *et al*., 2017). Introgression involves a backcrossing strategy with the ryegrass genome representing a greater proportion of the genomic material (Humphreys *et al*., 2003). Both methods allow for grasses with adequate yield and nutrition and increased hardiness and persistency providing ecosystem services

(Grogan *et al*., 2018; Macleod *et al*., 2013). *Festulolium* has been shown to out yield perennial ryegrass on marginal land for biomass production, illustrating the potential of hybrids under suboptimal growing conditions (Meehan *et al*., 2017). Though the production model in Ireland is based on grazed grass, silage is a significant constituent in beef and dairy production in Ireland and as grass growth is seasonal and its value varies within the year, the role of silage and conservation management is to "conserve" grass nutrients and meet demand in the production system during the winter feeding period (Shalloo *et al*., 2004; McEvoy *et al*., 2010; Mayne & O'Kiely, 2005). As it directly influences productivity, dry matter yield (DMY) has been the primary focus of breeding programmes in the improvement of perennial ryegrass varieties, including those destined for silage production (McDonagh *et al*., 2016). Persistency refers to stability of yield and this relationship with production has led persistency to be identifiable as an economically important trait. It is assessed both for Pasture Profit Index (PPI), the Irish index for economic value of cultivars and Value Cultivation and Use, for cultivars to be published on recommended lists (Grogan & Gilliland, 2011; O'Donovan *et al*., 2016). Persistency is widely quantified by ground score cover, a 1-9 scale, for referral to recommended lists (Grogan & Gilliland, 2011). Grass silage production involves mechanical cuts, with first and second cuts responsible for the greatest proportion of dry matter yield. Subsequent preservation under anaerobic conditions requires adequate levels water soluble carbohydrates (WSC) to ensure production of lactic acid and fermentation through microbial activity, and a relatively low buffering capacity (the acid required to lower the pH from 6 to 4) (Muck *et al*., 1988; Davies *et al*., 1998). Thus the nutrient profile of cultivars not only impacts the quality of animal nutrition but also the quality of silage production.

This trial was one part of a multinational collaboration coordinated by the Eucarpia Fodder Crops and Amenity Section *Festulolium* Working Group whereby the same cultivars were assessed across Europe in nine field trials and under contrasting growth conditions (Ghesquière *et al*., 2016). Partial results from this trial have been published previously, with data over a 3 year time period (Grogan *et al*., 2018). The aim of this study was to further investigate the agronomic performance of *Festulolium* hybrids, in comparison to their *Lolium* and *Festuca* parental species,

under conservation management over a five year time period under an Atlantic climate. Specifically, persistency measured by ground score cover, dry matter yield, and nutrient value of cultivars of the hybrid types were observed and whether the merits of each species are represented in their hybrids was investigated.

2.2 Materials and Methods

2.2.1 Site

This experiment was carried out in Athenry, Co. Galway, adjacent to the DAFM Recommended List trial site (53°18'N, 8°45'W), on a peaty loam soil. Meteorological data (Appendix 1) was obtained from a weather station, located on site.

2.2.2 Experimental design

Cultivars were donated, through EUCARPIA, from breeders across Europe; including six control species, representing *Lolium* and *Festuca* hybrid parental species, and 15 hybrids (Table 2.1). Trials were sown on 21st May, 2013 in field plots of 11.41m^2 utilising a randomised complete block (RCB) design with plots replicated in 3 blocks. Nitrogen (N) fertiliser ('High Sulphur CAN', 27.5 %N + 7 %S) was applied 6 times for a total annual rate of 310 kg N ha^{-1}. Based on annual soil analysis, no applications of phosphorus were used, and potassium was applied twice a year to a total of 400kg K ha^{-1}. The trial was cut using a Conservation/Silage protocol; incorporating two silage cuts and three after-grass cuts per year. The first two cuts were taken when the trial had reached stem extension/heading growth stage, while the later harvests were cut to schedule at vegetative growth stages

Table 2.1 Description of cultivars including the species/hybrid type (Lp = *Lolium perenne*, Lm = *Lolium multiflorum*, Fp = *Festuca pratense*, Fg = *Festuca glaucescens*, Fa = *Festuca arundinacea*), ploidy (2x = diploid, 4x= tetraploid, 6x = hexaploidy, all hybrids studied are tetraploid), hybridisation, associated breeder and publication.

Cultivar	Mode of hybridisation	Breeder	Reported in
Controls			
AberBite (*Lp4x*)		IBERS, UK	
AberMagic (*Lp2x*)	NA	IBERS, UK	
Caballo *(Lm4x*)	NA	DLF, Czech Repblic	
Podium (*Lm2x*)	NA	DLF, Czech Repblic	
Fure (*Fp2x*)	NA	Graminor Ltd., Norway	
Kora (*Fa6x*)	NA	DLF, Czech Repblic	
LpFp			
Fabel	Amphiploid	Graminor Ltd., Norway	Østrem *et al*. (2013)

FuRs0142	Amphiploid	Graminor Ltd., Norway	
Prior	Amphiploid	IBERS, UK	Lewis *et al*. (1973)
LmFa			
FuRs0352	Introgression	Graminor Ltd., Norway	
Bečva	Introgression	DLF, Czech Republic	Fojtik (1994)
Lofa	Introgression	DLF, Czech Republic	Fojtik (1994)
LmFp			
AberNiche	Introgression	IBERS, UK	
Agula	Amphiploid	Szelejewo, Poland	Zwierzykowski *et al*. (1998)
Felopa	Amphiploid	Szelejewo, Poland	Zwierzykowski *et al*. (1998)
Sulino	Amphiploid	Szelejewo, Poland	Zwierzykowski *et al*. (1998)
Hostyn	Introgression	DLF, Czech Republic	
Perseus	Amphiploid	DLF, Czech Republic	Houdek (2005)
Perun	Amphiploid	DLF, Czech Republic	Fojtik (1994)
Achilles	Amphiploid	DLF, Czech Republic	
LmFg			
Lueur	Amphiploid	INRA, France	Ghesquiere *et al*. (1996)

2.2.3 Measurements

Plots were harvested using a Haldrup plot harvester at cutting heights of ca six cm. Total plot yield was recorded and a subsample (circa 300 g) was oven-dried at 80 °C for 16 hours to determine DMY. Dried samples from the first silage cut in mid-May (over a three year period: 2015, 2016 and 2017), were analysed utilising established wet chemistry methods at Teagasc, Grange, where a subsample was dried at 40 °C for 48 h, milled through a 1 mm screen and analysed for in-vitro DM digestibility (DMD; Tilley and Terry, 1963), crude protein (CP; using a LECO FP-628 nitrogen analyser, AOAC 1990), water soluble carbohydrates (WSC; using the anthrone method; Thomas, 1997) and buffering capacity (BC; Playne and McDonald, 1966). Ground cover scores on a 1 to 9 scale, where 1 represented bare ground and 9 a fully grass-covered plot, were recorded by visual assessment at the end of each harvest year (procedure of visual scoring described in Lynch *et al*., 2015).

2.2.4 Statistical analysis

Analysis was carried out in R (R Core Team 2014; version 3.4.1). To investigate relationships between the measured parameters, type III analysis of variance (ANOVA) was implemented from the *car* package (Fox & Weisberg, 2019).

To further investigate the response of DMY of cultivars over time, a generalised least squares model was carried out, where correlation between years was modelled using a repeated measures covariance structure in the *nlme* package (Pinheiro *et al*., 2020) The response of total DMY was measured using the following models:

$$Y_{ijk} = \mu + T_i + R_j + C_k + (T_i \times C_k) + \epsilon_{ijk} \quad \text{(model 1)}$$

$$Y_{ijl} = \mu + T_i + R_j + H_l + (T_i \times H_l) + \epsilon_{ijl} \quad \text{(model 2)}$$

In model 1, Y_{ijk} denotes the total DMY for the year (T) i (2013 to 2018), replicate (R) within a block j (1-3), cultivar (C) k (1-21) and ϵ_{ijk} denotes the residual error. To investigate the response of DMY to hybrid type/pure species control model 2 was formulated where Y_{ijl} denotes the total DMY for the year i (2013 to 2018), replicate within a block j (1-3), hybrid type/pure species control l (1-10) and ϵ_{ijl} denotes the residual error.

A similar model was formulated to investigate the response of ground score to year:

$$P_{ijk} = \mu + T_i + R_j + C_k + (T_i \times C_k) + \epsilon_{ijk} \quad \text{(model 3)}$$

$$P_{ijl} = \mu + T_i + R_j + H_l + (T_i \times H_l) + \epsilon_{ijl} \quad \text{(model 4)}$$

In model 3, P_{ijk} denotes the total ground score for the year (T) i (2013 to 2018), replicate (R) within a block j (1-3), cultivar (C) k (1-21) and ϵ_{ijk} denotes the residual error. To investigate the response of DMY to hybrid type/pure species control (model 3). Model 4 was formulated where P_{ijk} denotes the total ground score for the year (T) i (2013 to 2018), replicate (R) within a block j (1-3), hybrid type/pure species

control (H) denoted by l (1-10) and ϵ_{ijk} denotes the residual error. A similar model to investigate the response of ground score to year. Post-hoc testing for significant differences between means was carried out using the *emmeans* package (Lenth, 2020). Assumptions for analysis were checked and met.

2.3 Results

2.3.1 Persistency expressed as ground cover

Overall the trial had adequate ground cover (GC) for the first four years after field establishment, however ground cover deteriorated in year 5 (Figure 2.1). The mean ground cover for all years varied significantly from one another, besides the years 2015 and 2016, which were found to not be significantly different (Supplementary Table 2.2). The lowest ground cover scores were observed in 2018, and the highest ground score cover scores were found in 2014. Year, cultivar, hybrid type and block were all found to have a significant effect on ground cover score, and there were significant interactions between year and cultivar, and between year and hybrid type (Appendix 2; Appendix 3).

No significant differences were observed within hybrid types for the first three years, however within the *Lolium multiflorum* x *Festuca pratensis* hybrid type, Sulino was found to vary significantly from AberNiche, Perseus and Achilles for mean ground score cover in 2018 (Appendix 6). Within *L. multiflorum* x *F. arundinacea* hybrids, Lofa and FuRs0352 varied significantly in 2017 and 2018, and FuRs0352 and Bečva exhibited significant differences from one another in 2018 (Appendix 6). There were no significant differences observed between hybrids within *L. perenne* x *F. pratensis* for ground score cover within any experimental year (Appendix 6). Of the pure species controls, both tetraploid and diploid *L. perenne* controls exhibited the highest ground cover of over time, followed by pure *F. pratensis* (Appendix 4).

From the *Festulolium* hybrids the best ground cover scores were obtained from *L. perenne x F. pratensis* hybrids, with the cultivar Fabel exhibiting the highest mean ground score cover in comparison to all other cultivars (Appendix 4). The poorest ground cover score from the beginning to the end of the trial was obtained by the *Festulolium* hybrid Lueur of the *L. multiflorum x F. arundinacea* var. *glaucescens* type, while FuRs0352, a *L. multiflorum x F. arundinacea* hybrid, was the cultivar

with the lowest over all mean ground score cover (Appendix 6). Three statistically significant groups are present from mean ground score cover across 5 years, with tetraploid *L. perenne* and hybrid type *L. perenne x F. pratensis* belonging to the highest performing grouping, while the remaining hybrids and *F. pratensis* allocated in the group with the lowest ground cover values (Appendix 6). The remaining pure species controls overlap with these groups and share a group among themselves (Appendix 6).

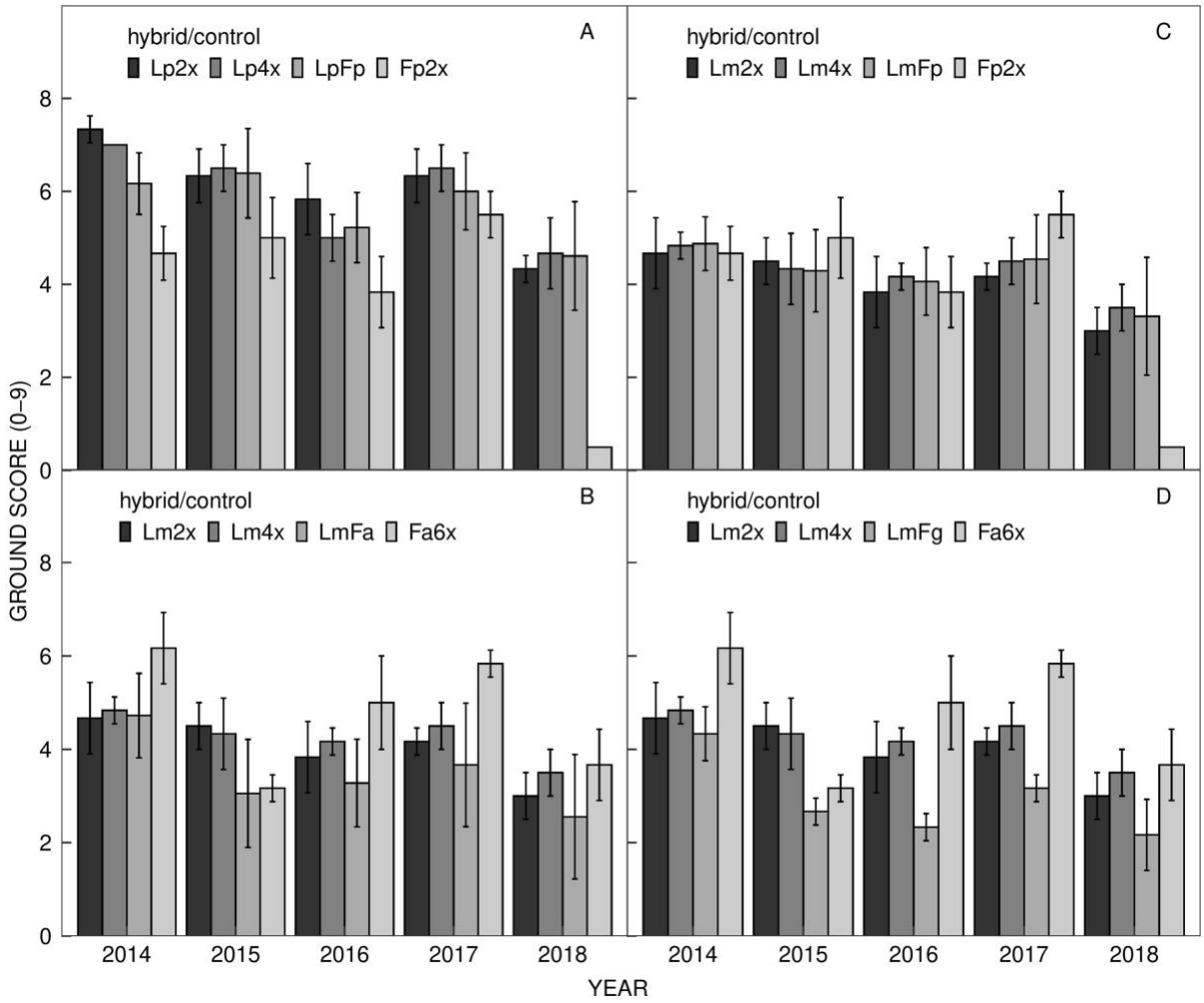


Figure 2.1. Trends in the mean ground cover score for each of the *Festulolium* hybrid types (A) *L. perenne x F. pratensis*, (B) *L. multiflorum x F. arundinacea*, (C) *L. multiflorum x F. pratensis*, (D) *L. multiflorum x F. arundinacea* var. *glaucescens,* alongside their respective control species over a five year period.

2.3.2 Yield

The first cut of silage accounted for, on average, 46.6 % of total DMY, while the second cut accounted for an average of 26.2 % of total DMY. The remaining

percentage of total DMY was comprised of the subsequent three cuts. Similar trends in total DMY production were observed across hybrid types, with the establishment year exhibiting lower yields, increasing in 2015, declining in 2016, recovering in 2017 and then finally declining in 2018 (Figure 2.2). This was reflected when comparing the mean yield for each year; the mean yields for 2014 and 2016 did not differ significantly, likewise, the mean yield for 2018 and 2016 did not exhibit significant differences (Appendix 7). Year and cultivar had a significant effect on DMY, with the interaction between the variables not found to be significant. Statistically significant differences in yield were also observed amongst hybrid types with a significant interaction with year. *L. multiflorum* x *F. pratensis* hybrids Hostyn and Perseus outperformed other cultivars of the same hybrid type in yield, though there was no statistical significance between the cultivars for each year. Similarly, the *L. multiflorum x F. arundinacea* hybrid Bečva performed well across years in comparison to *Lolium* controls, however mean DMY for the cultivar declined significantly from 2015 to 2016, with 2018 yields also found to be significantly different form 2015. Lower yielding varieties and controls exhibited no such significant decrease in yield. Overall, varieties which exhibited lower yield in earlier years remained consistent, while higher yielding varieties declined in yield in 2018.

Two distinct groups can be observed when the mean yield of hybrid types are compared; *L. mutiflorum* tetraploid control was allocated to the highest yielding group, along with the *Festulolium* hybrids *L. multiflorum* x *F. pratensis* and *L. multiflorum* x *Festuca* (Table 2.1). *F. pratensis* diploid control was categorised into the second, lower yielding group, with the remaining hybrids/controls overlapping with either group.

Table 2.2. The mean total dry matter yield (DMY), per cultivar for each year and the overall mean for a five year period, including mean and standard deviation (SD) for yield per hybrid type for each year and five year mean.

Cultivar	**2014 DMY (g m^2)**	**2015 DMY (g m^2)**	**2016 DMY (g m^2)**	**2017 DMY (g m^2)**	**2018 DMY (g m^2)**	**Mean DMY (g m^2)**	**SD DMY (g m^2)**
Controls							
AberBite (*Lp4x*)	1443	1764	1306	1592	1140	1449[a,b]	246
AberMagic (*Lp2x*)	1453	1810	1337	1561	1355	1503[a,b]	226

Caballo (L*m4x*)	1554	1004	1367	1692	1391	1607[b]	262
Podium (*Lm2x*)	1209	1918	1411	1585	1488	1522[a,b]	263
Fure (*Fp2x*)	1106	1334	1165	1310	1290	1241[a]	121
Kora (*Fa6x*)	1273	1523	1348	1315	1425	1377[a,b]	127
LpFp							
Fabel	1504	1650	1282	15.74	1307	1464	221
FuRs0142	1457	1611	1377	14.84	1320	1450	220
Prior	1332	1612	1250	1543	1305	1410	180
Mean	1431	1626	1303	1534	1311	1441[a,b]	-
SD	181	193	158	1.72	121	-	204
LmFa							
FuRs0352	1267	1698	1250	1543	1304	1396	240
Bečva	1547	1949	1486	1744	1237	1593	380
Lofa	1496	2036	1436	1823	1261	1610	299
Mean	1437	1894	1398	1636	1300	1533[b]	-
SD	163	183	132	132	202	-	285
LmFp							
AberNiche	1383	1864	1402	1499	1105	1451	279
Agula	1455	1704	1285	1453	1251	1430	243
Felopa	1479	1775	1398	1582	1224	1492	178
Sulino	1426	1704	1362	1480	1332	1461	252
Hostyn	1534	1925	1577	1869	1396	1660	252
Perseus	1619	1936	1515	1872	1475	1683	217
Perun	1483	1753	1438	1770	1368	1544	246
Achilles	1522	1882	1426	1644	1368	1568	234
Mean	1488	1818	1425	1646	1303	1536[b]	-
SD	112	201	129	230	154	-	246
LmFg							
Lueur	1245	1538	1263	1508	1294	1370[a,b]	153

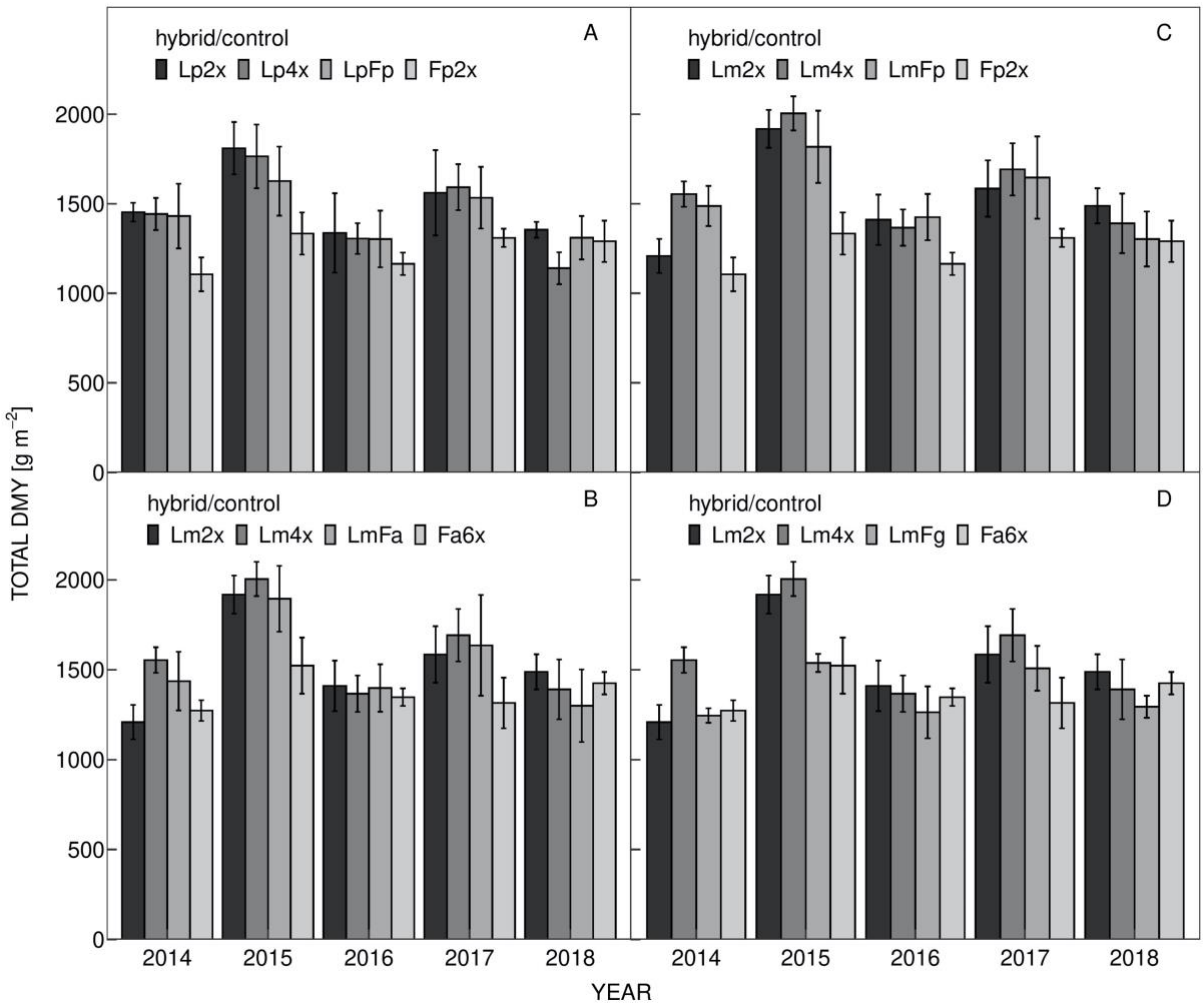


Figure 2.2. Trends in the mean dry matter yield (DMY) for each of the *Festulolium* hybrids for each of the hybrid types (A) *L. perenne x F. pratensis*, (B) *L. multiflorum x F. arundinacea*, (C) *L. multiflorum x F. pratensis*, (D) *L. multiflorum x F. arundinacea* var. *glaucescens,* alongside their respective control species over a five year period.

2.3.3 Nutritional profile in first cut silage

Year, cultivar and hybrid type all had a highly significant effect ($p<0.001$) on soluble carbohydrates, buffering capacity, protein content, dry matter digestibility of the first silage cuts over a three year period (2015 to 2017) (Appendix 2; Appendix 3). There were no significant interactions between year and cultivar or year and hybrid type for any for the variables, with the exception of a significant interaction observed in WSC between year and cultivar.

DMD steadily declined over the three year period, across all hybrid types and pure species controls (Figure. 2.3). The hybrid type exhibiting the highest mean DMD in 2015, 2016, 2017 was observed to be *L. perenne x F. pratensis* (mean = 713.96 g

kg (d.m.)), standard deviation = 38.30 g kg (d.m.)) and the pure species control, and highest mean over all was the tetraploid *L. perenne* (mean = 760.67 g kg (d.m.) standard deviation = 46.24 g kg (d.m.)). *L. multiflorum x. F. arundinacea* var. *glaucescens* was the hybrid type which exhibited the lowest mean DMD (mean = 689.96 g kg (d.m.), standard deviation = 46.90 g kg (d.m.)), and *F. arundinacea* was the control with lowest DMD over all three years (mean = 696.70 g kg (d.m.), standard deviation = 38.23 g kg (d.m.)).

Values for WSC declined in 2016 and recovered in 2017, and similar to the trends in DMD, trends were largely conserved across all hybrid types and pure species controls. Higher WSC was observed in *L. multiflorum* controls with *L. perenne* and hybrids exhibiting intermediate to lower levels of WSC, with *Festuca* pure species controls the lowest (Figure 2.3). The hybrid type which had the highest observed mean for WSC was *L. multiflorum x F. arundinacea* (mean = 241.67 g kg (d.m.), standard deviation = 59.36 g kg (d.m.)), with tetraploid *L. multiflorum* the pure species control with the highest mean WSC (mean = 280.82 g kg (d.m.), standard deviation = 45.59 g kg (d.m.)). The hybrid type exhibiting the lowest mean in 2015, 2016, 2017 was observed to be *L. multiflorum x F. arundinacea* var. *glaucescens* (mean = 152.22 g kg (d.m.), standard deviation = 14.25 g kg (d.m.)) and the control *F. pratensis* (mean = 129.92 g kg (d.m.), standard deviation = 18.38 g kg (d.m.)).

Higher BC was observed in *Festuca* controls and *Festulolium* hybrids than in *Lolium* controls (a lower buffering capacity is desirable in relation to ensilibility and silage preservation) be *L. multiflorum x F. arundinacea* var. *glaucescens* (mean = 361.44 g kg (d.m.), standard deviation = 62.99 g kg (d.m.)) and pure species control *F. arundinacea* the control *F. pratensis* (mean = 394.28 g kg (d.m.), standard deviation = 44.13 g kg (d.m.)). A decrease in BC was observed over time (Figure 2.3), with some exceptions *F. arundinacea*, gaining BC from 2016 to 2017 (mean = 315.32 g kg (d.m.); 383.64 g kg (d.m.), standard deviation = 25.52 g kg (d.m.), 96.33 g kg (d.m.), respectively). Lowest level of buffering capacity was observed in the *L. multiflorum* pure species controls and intermediate levels in *L. multiflorum* and *L. perenne* hybrids (Figure 2.3).

Trends in CP were similar to that of BC, in that there was a steep decline between 2015 and 2016 with a more gradual downward gradient between 2016 and 2017, and though there was generally consistency in this trend among pure species controls and hybrid types there were some exceptions. The pure species *F. arundinacea* increased in CP between 2016 to 2017 (mean = 88.28 g kg (d.m.); 94.26 g kg (d.m.), standard deviation = 3.91 g kg (d.m.); 2.2 g kg (d.m.), respectively) with more minimal increases for the same period in *L. multiflorum x F. arundinacea* (mean = 68.99 g kg (d.m.); 71.20 g kg (d.m.), standard deviation = 17.37 g kg (d.m.); 96.33 g kg (d.m.), respectively) and *L. perenne x F. pratensis* (mean = 66.36 g kg (d.m.); 71.06 g kg (d.m.), standard deviation = 21.08 g kg (d.m.), 0.79 g kg (d.m.), respectively) (Fig. 3). Consistently across all years, *F. pratensis* (mean = 125.63 g kg (d.m.), standard deviation = 25.23 g kg (d.m.)) exhibited the highest levels of CP. The lowest mean CP across three years was observed in the tetraploid *L. multiflorum* (mean = 77.92 g kg (d.m.), standard deviation = 13.17 g kg (d.m.)), followed by the hybrid *L. multiflorum x F. arundinacea* (mean = 79.85 g kg (d.m.), standard deviation = 21.43 g kg (d.m.)).

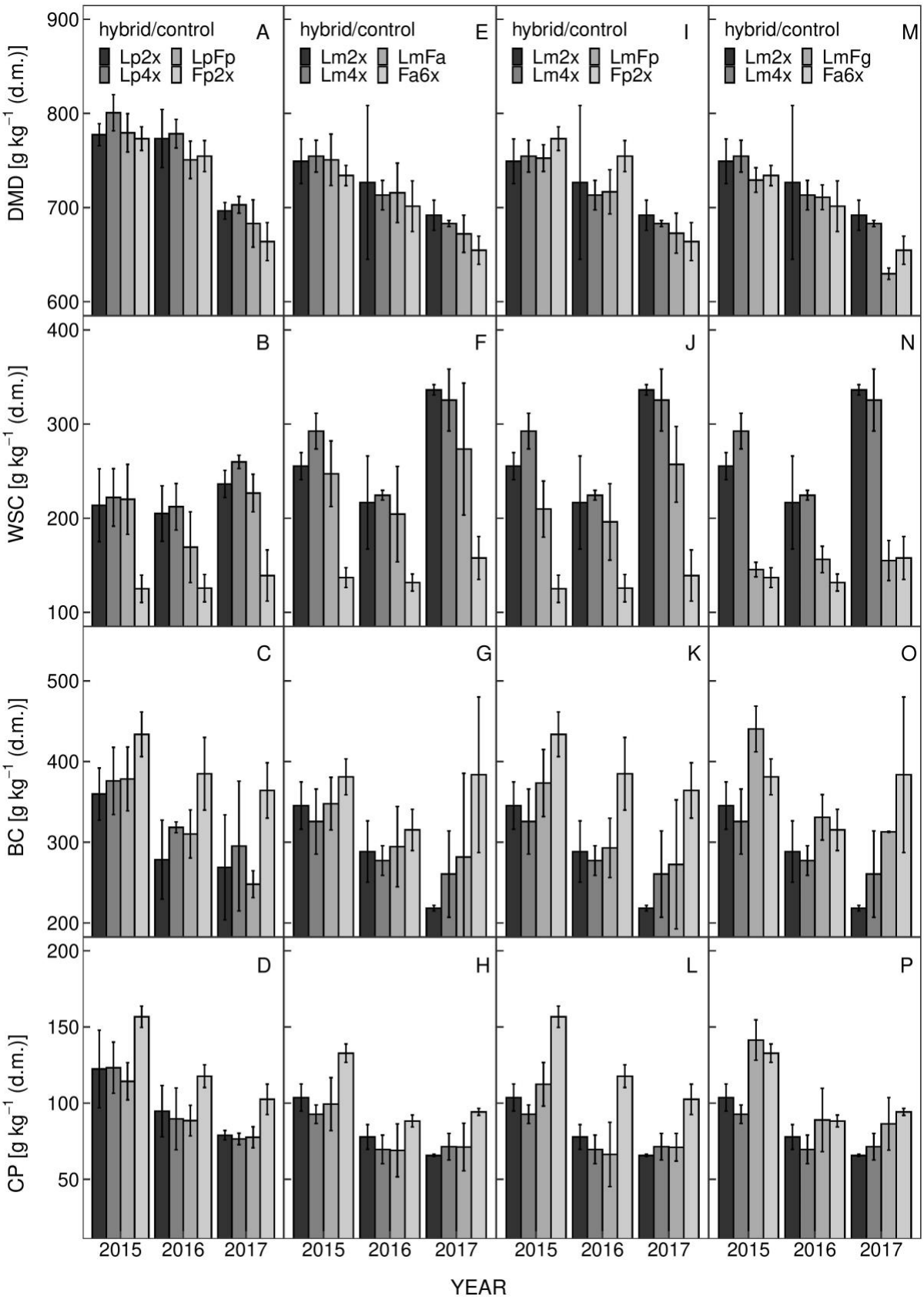


Figure 2.3. Quality parameters over three years, dry matter digestibly (DMD); water soluble carbohydrates (WSC) buffering capacity (BC); crude protein (CP); for each of the *Festulolium* hybrids for each of the hybrid types (A, B, C, D) *L. perenne x F. pratensis*, (E,

F, G, H) *L. multiflorum x F. arundinacea*, (I, J, K, L) *L. multiflorum x F. pratensis*, (M, N, O, P) *L. multiflorum x F. arundinacea* var*. glaucescens,,* alongside their respective control species over a three year period.

2.4 Discussion

Festulolium hybrids are of interest for forage production due to their complementary characteristics, with *Lolium* exhibiting greater quality attributes and *Festuca* possessing superior resilience to abiotic stress. We carried out a five year field experiment to evaluate dry matter yield, ground cover and persistency, and nutritional parameters of *Festulolium* hybrids under a conservation management in a mild Atlantic climate.

Persistency is a complex trait controlled by many interacting genetic and environmental factors, and it has been identified as a trait for selection in perennial ryegrass breeding programmes due its economic importance associated with sward stability and renewal (Conaghan & Casler, 2011). The trends in persistency were conserved among hybrid types, and they most aligned more closely to the *Lolium* pure species controls. Values for ground score, however demonstrated comparability between *L. multiflorum* controls, hybrid types, and *Festuca* controls when five year means were grouped. *L. perenne* pure species controls exhibited greater persistency, which was expected (Wilkins & Humphreys, 2003). The hybrid *L. perenne x F. pratensis* also displayed ground score similar to that of its *Lolium* component*. L. multiflorum* controls were expected to be the least persistent, however they were they were found to have intermediate persistency (Wilkins & Humphreys 2003). There was some deviation within hybrid types, with some cultivars deviating significantly from hybrid groups in later years of the trial. The *Festulolium* hybrid Spring Green cv. demonstrated superior ground cover in colder climatic in a multisite study in the USA, while performing poorer in warmer regions in comparison to the *Festulolium* hybrids from which it was derived (Casler *et al*., 2002) This indicates the variation the wide adaptability of *Festulolium* in terms of persistency. *Festuca* pure species control exhibited poorer persistency based on ground score. Due to the resilience to abiotic stress observed in *Festuca*, it was expected for these cultivars to exhibit enhanced persistency. However, ground score cover is a measure of abundance and does not necessarily reflect productivity

as demonstrated by the cultivar Fure, the *F. pratensis* pure species control. Although ground score decreases, DMY increases after an extended period, perhaps due to the broad leaf morphology of *Fescue* species in comparison to *Lolium* species (Gymer & Whittington, 1972). Despite the presence of open ground indicated by the low ground score, the larger leaf area and subsequent larger volume of biomass may account for the disparity between ground score as a measure of persistency and the observed yield. Though ground score cover is the widely accepted measurement of persistency for national list in Ireland, it is not without limitations. Ground score measures the abundance of a species within a sward rather than the biomass of the surviving plants, failing to account for yield stability. Methods have been proposed such as reporting the relationship between short term and medium term yields to describe persistency (Dodd *et al*., 2018).

In addition to ground cover scores, the change in dry matter yield over time is also an important component of persistency. The results from this trial for DMY demonstrate *Festulolium* hybrids perform well in terms of productivity and in some cases out yield *Lolium* controls. Similarly, Meehan *et al*. (2017) found that *Festulolium* hybrids along with *F. arundinacea* were capable of producing greater DMY than *L. perenne* under suboptimal conditions, over a two year period for biomass production. Establishment years tended to exhibit poorer yields in grasses as well as sharp declines in DMY in 2016 and 2018. Though during these years higher yielding hybrid types and controls declined, they still exhibited higher DMY values than more stable yielding varieties. Yield stability is more prevalent in lower yielding hybrids and similar results were found in two Swedish regions where *Festulolium* varieties which yielded higher declined more quickly, while *Festulolium* varieties with lower yields remained stable (Halling, 2012). In a Lithuanian climate *Festulolium* exhibited superior yields to both *Festuca* and *Lolium* with *L. perenne* displaying much lower yields (Lemežienė, 2004). The decline in DMY in 2016 and 2018 observed in this study was likely due to low rainfall during the slate spring and early summer months (April 2016 = 49.2mm, May 2016 = 56.7mm, May 2018 = 62.5mm, June 2018 = 25.2mm). Based on historical data, low summer rainfall is generally not a limiting factor in Irish climatic conditions, except when values for total monthly rainfall reached lows of between 20-60mm during the late spring to

summer period, as observed in both 2016 and 2018 (Hurtado-Uria *et al*., 2014) the climatic influence causing variability among hybrids over a single year was observed in a multisite study including that described here (Ghesquière *et al*., 2016). In terms of yield in this study, the hybrids yields appear to more closely follow their *Lolium* component than their *Festuca* component. This was also observed when foliar characteristic and DMY were observed in *Festulolium* hybrids and compared to that of *Lolium* controls (Humpreys *et al*., 2014; Muhandiram *et al*., 2020). This may be due to breeding strategy, genome instability after successive generations, or disparity between integration of source genomes, such as ampliploid or introgressed hybrids. We did observe some variability among hybrid cultivars, where some cultivars exhibited comparable yields to *Festuca* while others out yield *Lolium*, though these cases were not found to be significant.

In general, *Lolium* has been observed to possess higher digestibility, therefore, in this study we expected that *Lolium* controls would exhibit relatively higher levels of quality, followed by *Festulolium* hybrids and finally, *Festuca* controls with the lowest quality values (Banzhaf and Boberfeld, 2005; Wilkins and Humphreys, 2003). We observed that all pure species cultivars and hybrids were affected equally by different experimental years which points towards a good silage quality under marginal conditions. *Festulolium* cultivars exhibited intermediate levels for both WSC and BC parameters, indicating they would be suitable for silage production. Similary, no significant differences between *Festulolium* hybrids *Lolium perenne* controls and was observed in a previous study when digestibility of hybrids was investigated (Kamau *et al*., 2020) The *F. arundinacea* cultivar Kora had the highest levels of CP across all years. This was replicated in a Danish study where Kora was found to have higher CP content in the first cut of silage in comparison to *Festulolium* hybrid (Solati *et al*., 2018) with CP declining to similar levels in both cultivars in the autumn. This is likely due to maturity, as the range of heading dates will affect the as the plants move from vegetative to reproductive phase, more lignin the nutrient profile of the plant becomes altered.

2.5 Conclusion

Overall, *Festulolium* hybrids performed at a very similar level as *Lolium* species in terms of persistency, yield and ensilibility in this five year study. This makes

Festulolium hybrids very suitable to be grown as a silage crop under a mild Atlantic climate, even under marginal conditions.

Chapter 3

Detecting compositional changes in perennial ryegrass populations utilising a low cost, multiplex genotyping tool, GenoGrass

Abstract

The use of molecular markers in perennial ryegrass populations has the potential to accelerate the genetic gain in many economically important traits, particularly in persistency. The cost and technical challenges associated with employing next generation sequencing often prohibits its use for many applications in perennial ryegrass (PRG) improvement. Genotyping by sequencing (GBS) approaches are characterised by the ability to survey thousands of sites, however, some applications may not require such wide genome coverage, but would benefit from a GBS-based system that allowed the low cost analysis of thousands of samples at hundreds of loci. The design of a low-cost high throughput genotyping platform was undertaken to mitigate these difficulties. A total of 192 sites were selected for primer design, 54 of which were identified from previous research which had high predictive ability for heading date and crown rust resistance, while the remaining sites were selected for their even distribution across the PRG genome. This assay was then deployed in a number of populations with varying levels of complexity. Population types include individuals from breeding populations, ecotypes as well as mixed pooled populations with varying levels complexity. The efficiency of the assay was tested in identifying population structure and diversity across multiple population types. We found the assay to be successful in discerning population structure in breeding populations, ecotypes, cultivars and complex mixtures. An adequate depth of coverage was achieved, however an overrepresentation of some sites likely influenced the overall evenness of coverage. Reduced representation assays are suitable for the characterisation of PRG populations, however improvement of elements of the assay are necessary for future use.

Key words: Genotyping by sequencing, GT-Seq, *Lolium perenne*, SNP

3.1 Introduction

The prevalence of perennial ryegrass (*Lolium perenne*) in temperate grazing systems highlights the importance of the species as a forage crop (Wilkins & Humphreys, 2003). In Ireland, single species swards of mixed cultivars are the most commonly sown with perennial ryegrass (PRG) accounting for 95% of the grass seed sold (DAFM, 2021). PRG serves as a cheap source of feed for bovine agriculture, reducing costs and increasing profits (O'Donovan *et al*., 2011). The use of molecular markers, often regions which are inherited with a trait of interest, has the potential to accelerate the improvement of economically important traits (Hayes *et al*., 2013). Molecular markers have been used to improve traits in PRG such as maturity, disease resistance, yield, and quality; however traits such as persistency have not undergone the same improvement (Arojju *et al*., 2016; McDonagh *et al*., 2016; Arojju *et al*., 2018). Persistency refers to yield stability over time, and relationship with production has led persistency to be identifiable as an economically important trait (O'Donovan *et al*., 2017). Investigating genetic diversity in single species swards can be used to observe changes in population structure over time (Verwimp *et al*., 2018). This may allow for the identification of more persistent genotypes, and subsequently, for the identification molecular markers for persistency.

As an obligate outbreeding species, PRG populations are genetically diverse with high levels of geneflow within and among populations (McGrath *et al.*, 2007; Diekmann *et al.*, 2012). As a result, PRG cultivars consist of heterogenous populations comprised of heterozygous individuals. PRG also possesses a large genome (~2.2 Gbp) with high levels of repetitive sequence content (76%) (Byrne *et al*., 2015). These aspects of PRG genomics can result in technical challenges in the context of sequencing experiments. Improvements in high throughput next generation sequencing (NGS) technologies have greatly increased the quantity of sequence data used in studies over the last decade and has contributed to the improvement of genetic resources and SNP discovery in PRG (Byrne *et al*., 2016; Veekman *et al*., 2016). However, a limitation of data generated by NGS is the ubiquity of a non-negligible error rate, often causing ambiguity in mapping and distinguishing between accurate and erroneous variant calls. Strategies to mitigate

this usually involve bioinformatic pipelines that employ filtering strategies at multiple stages of processing, based on likelihoods of error. However approaches such as these can be problematic in high diversity species such as PRG, where rare alleles offer greater informative and discriminatory power in population differentiation, but are likely eliminated in the interest of excluding spurious variants (Linck and Battey *et al.*, 2019). By increasing the number of reads per locus, theoretically, SNPs can be more reliably identified, though evenness of coverage across loci is not guaranteed (Sims *et al*., 2014). Furthermore, the associated expense of increasing the number of generated reads can be cost-prohibitive when sequencing a large number of samples, resulting in a trade-off between depth and breadth of coverage as well as sample number.

Genotyping by sequencing (GBS), as described by Elshire *et al*. (2011), utilises restriction enzymes to reduce genome complexity and the resulting libraries are sequenced using NGS platform. This enzymatic approach has the capacity to generate 100,000s of SNPs and lends itself to applications which require high levels of genome coverage and the generation of new markers (Pembleton *et al*., 2016; Arojju *et al*., 2018). Though this technique has been utilised in a broad range of studies, for many applications this level of genome coverage is unnecessary. A relatively low number of SNPs (~50) have been shown to be responsible for a significant portion of heritability in agriculturally important traits in PRG (Byrne *et al.,* 2017; Arojju *et al*., 2018). Reducing the number of surveyed sites, and increasing depth of coverage can reduce costs on a per sample basis, while mitigating for high diversity crop species. A number of both commercially available and freely available low cost, targeted resequencing technologies are available which have been utilised in crop species, including Ampliseq, SNP-seq, Single Primer Extension Technology (SPET) and Genotyping in Thousands by sequencing (GT-seq) (Campbell *et al*., 2015; Barchi *et al*., 2019; Ogiso-Tanaka *et al*., 2019; Zhang *et al.*, 2020). GT-seq is an amplicon based system of genome complexity reduction. It utilises a dual barcoding system and subsequent high-throughput NGS to cheaply and efficiently sequence a reduced panel of target loci in a large number of samples. This approach is particularly useful in targets for which there is *a priori* knowledge of sequences. As well as employing reduced representation assays, pooling samples

can reduce costs, improve efficiency and has been shown to accurately genotype perennial ryegrass populations (Byrne *et al*., 2013; Verwimp *et al*., 2018).

The aim of this investigation was to develop a low cost, high throughput method for genotyping platform “GenoGrass”, based on GT-seq. To do so, we utilised an array of target SNPs which have been identified from previous research, as well SNPs which are evenly distributed throughout the PRG genome. These targets were deployed in a number of populations with varying levels of complexity, in both samples from individuals and pools of individuals, to test the efficiency of the assay in identifying population structure and diversity across multiple population types. We assessed the efficiency of the assay in terms of depth, evenness of coverage, accuracy of targets as well as the ability to detect population structure.

3.2 Methods

3.2.1 Plant material

This investigation used DNA from plant material derived from three previous investigations, contributing to a total of 665 samples (Table 3.1). Table 3.1 describes the three populations included in this investigation.

A fourth population was subsampled from a F2 biomass population This mapping population was not included in this investigation, however the population was sequenced and is included in the total read counts (Appendix 8).

3.2.1.1 Breeding population

A total of 473 individuals were subsampled from six synthetic cultivars and eight full-sib families (n=10-45). These individuals were subsampled from the Teagasc breeding programme and were used previously in a genomic selection study to assess the predictive ability markers for heading date and crown rust resistance (Byrne *et al*., 2017; Arojju *et al*., 2018).

3.2.1.2 Ecotype

The second population was derived from a diversity study, investigating the genetic variation in an Irish genetic resource collection of PRG (Barth *et al.*, 2017). A subset of these individuals, 140 individuals from 11 accessions, were selected and included in this study (n=9-29).

3.2.1.3 Artificial mixture

The final population is comprised of an artificial mixture population generated as part of an unpublished previous study. All sampling, library preparation and preliminary data analysis was conducted by Dr Aude Perdereau. The populations were generated by germinating PRG seed in glasshouse conditions with 1g of seed per pot, in 9x9cm pots filled with soil. Six F2 diploid PRG families (A, B, C, D, E, F) were sown individually in pots, replicated by four, and in various ratios (Table 3.2). Ratios were chosen to investigate the ability of the assay to subtle changes in mixtures of varying complexity. Ratios of seed for each mixture component was measured by weight. Two weeks post-sowing, 1cm of plant material was harvested from each pot. Plant material was manually ground using liquid nitrogen and DNA was isolated using cetyl trimethylammonium bromide (CTAB), as described by Doyle (1991).

Table 3.1. Description of populations that were genotyped with the GenoGrass assay (superscripts denote the following breeders; [1]= IBERS; [2]= Teagasc [3]= AFBI; [4]= Euro Grass; [5]= DLF-Trifolium; [6]= NPZ; [7]=RvP; [8]= Norddeutsche; [9]= Cebeco; [10]=Zelder)

Group	Sample number	Sample type	Publication	Population type	ID	Cultivar/Origin/
Breeding population	473	Individuals	Bryne *et al*., 2016, Arojju *et al*., 2018	Cultivar	G1	Aberstar[1]
					G4	Genesis[2]
					G7	Tyrella[3]
					G8	Malambo[4]
					G9	Boyne[5]
					G10	Glenroyal[2]
				Full sib	G11	Pastour[5]
					G12	Soloman[2]
					G13	Jumbo[5] X Tyron
					G14	(Donard[3] X Morg
						(Donard[3] X Corb
					G15	Profit[9] X Hercule
					G16	AberAvon[1]
					G17	Tyrconnell[2]
					G18	AberSilo[1]
Ecotype	140	Individuals	McGrath *et al*., 2007, McGrath *et al*., 2010, Barth *et al*., 2017	Ecotype	COR	Cork, Ireland
					GAL	Galway, Ireland
					KER	Kerry, Ireland
					LAO	Laois, Ireland
					MAY	Mayo, Ireland
					OFF	Offaly , Ireland
					ROS	Roscommon, Ire
					TIP	Tipperary, Irelan
					WEX	Wexford, Ireland
				Cultivar	Fennema	Fennema[6]
					Porstewart	Portstewart[3]
Artificial mixtures	52	Pools	Unpublished	F2	A	Pastour[5]
					B	Solomon[2]
					C	Tyrconnell[2]
					D	Profit[9] X Hercule
					E	Jumbo[5] X Tyron
					F	(Donard[3] X Morg
						(Donard[3] X Corb

Table 3.2. Artificial mixture population description and mixture ratios

Pure families		**Two family mixtures**		**Multiple family mixtures**	
Family	*Ratio*	*Family*	*Ratio*	*Family*	*Ratio*
A	100	A + B	50:50	A + B + C + D + E + F	Equal ratio
B	100	A + B	75:25	A + B + C + D + E + F	10:10:40:25:5:10
C	100	A + B	87.5:12.5	A + B + C	45:45:10
D	100	A + B	96.5:3.5		
E	100				
F	100				

3.2.2 Target design

Target selection for primer design is a key step in this assay in relation to the technical efficiency of the multiplex PCR steps, to avoid unwanted annealing and artefacts. It is also an important consideration from the perspective of generating genotypes, as the value of markers relies on informative SNPs.

Initially, 212 target amplicons were designed. A total of 79 of these amplicons were selected from Genomic Selection experiments carried out on the aforementioned population. A total of 69 target loci were selected, as the correspond to SNPs which were identified as having high predictive ability for heading date (HD) and crown rust resistance (CR), identified from GBS data and ranked using variable importance measures (Byrne *et al*., 2016; Arojju *et al* ., 2018). The remaining 123 targets were selected based on the occurrence of SNPs which were in anchored scaffolds at relatively even intervals across the seven perennial ryegrass linkage groups (LG) as defined by the genome zipper (Byrne *et al*., 2015). Scaffolds in the genome zipper were anchored based on the syntenic relationship between PRG and other members of the grass family (*Poacae*). Scaffolds confidently anchored using the genome zipper were selected and regions with high levels of variation were avoided to mitigate primer failure. Primer pairs were successfully designed for 85% of amplicon targets identified initially (LGC, Berlin). Reasons for failure during primer design included presence of GC rich regions and an overabundance of polymorphisms present in the region of interest. Targets which failed during the primer design phase, including those derived from the GS datasets, were replaced with targets containing randomly selected SNPs that were anchored to the genome

zipper. Length of target sequences post primer design ranged between 150 and 221 bp. The final array consisted of 192 target amplicons (Appendix 8). In total, 177 targets were anchored to the genome zipper, 153 with stringent parameters for anchoring and 24 with non-stringent. Twenty four out of 36 targets from the heading date data set and 30 out of 33 targets from the crown rust data set were anchored. The remaining 123 anchored amplicon targets were designed to give relatively even coverage of the perennial ryegrass genome based on their position on the perennial genome zipper.

3.2.3 Genotyping protocol

The GenoGrass protocol is based on GT-seq, developed by Campbell *et al.* (2015). The protocol involves a two-fold multiplex PCR; the first reaction amplifies targets and attaches sequence primers, while the function of the second reaction is to attach barcodes to the amplified products from the first PCR. A pooling and normalisation step of PCR product is carried out, and all amplicons are sequenced using Illumina NextSeq. Two rounds of sequencing was performed to raise representation of underrepresented targets. The design of primers to capture targets, preparation of libraries and sequencing was performed by LGC (Berlin, Germany).

3.2.4 Data analysis

GenoGrass samples were sequenced using paired end reads. Reads were demultiplexed using bclfastq v. 2.20 (Ilumina, San Diego, CA, USA), allowing a maximum if 2 mismatches. Reads were clipped of adaptor sequences and trimmed for poor quality bases. Paired reads were aligned to the PRG reference genome using bwa v. 0.7.17 (Li & Durbin, 2010). Depth of coverage for each target site was using the SAMtools v. 1.2 (Li *et al.,* 2009), coverage function, where a bed file containing targets positions was used to generate the number of paired end reads per site. Variants were called using vcftools v. 0.1.16 mpileup (Danecek *et al*., 2011). Variant sites with a genotype quality of at least 30, a read depth of more than 10 and missingness of less than 50% were retained. The read depth and missingness thresholds were reasonably relaxed to capture as many target sites as possible. Variant sites with minor allele frequency of less than 0.1 were also excluded from the final set. All statistical analysis was performed in R. The adegenet

v. 2.1.3 package in R (Jombart, 2008; Jombart *et al*., 2010) was used to perform Discriminant Analysis of Principal Components (DAPC) analysis on the breeding population and ecotype populations, utilising the genotypes of individuals and visualising the results. The posterior probability based on the retained discriminant functions analysis included membership assignment probability, and was visualised. The number of principal components (PC) to be retained was determined using the dudi.pca function from the ade4 v. 1.7.15 package (Dray and Dufour, 2007). The artificial mixtures pooled population genotyped with the GenoGrass assay was analysed using principal component analysis (PCA) based on calculated allele frequency of each variant of each pooled sample. Ward's minimum variance method was used to perform hierarchical cluster analysis. The population was also genotyped using GBS, with a double enzyme digest. This was aligned using bwa, with the variant calling performed with GATK's HaplotypeCaller v. 4.1.0 (McKenna *et al.,* 2010) (carried out by Dr Perdereau). Resulting variants were included in this study for comparison and were filtered and analysed as described for the GenoGrass populations.

3.3 Results

3.3.1 Read alignment, depth and distribution

Sequencing produced 178 million paired end reads in 768 samples, including the three experimental populations and an unanalysed population (see methods). Two rounds of sequencing was performed due underrepresentation of some target regions. The total alignment rate of reads was evaluated after reads were demultiplexed, trimmed, paired and aligned. The mean alignment rate of properly paired reads was 86.02%, the lowest rate of alignment was 0%, with a high of 99.42% (Appendix 9). A group of 44 samples from the breeding population had an alignment rate of 0%. Upon investigation, these sample originated from the same locations within a 96 well plate and may have been a result of technical error and these values were removed from summary statistics as they do not accurately reflect the efficiency of the assay.

The distribution of properly paired reads per sample per target region (RD) across the entire data set of 192 targets, is represented in 665 samples. Summary statistics indicate a high level of variability in target representation efficiency and accuracy

(Table 3.3). The values described here correspond to paired end reads, with the average number for target region reported. The mean of 180.03 RD, across all populations, is well above the desired coverage. However extreme values due to overrepresented loci are inflating mean values, as indicated by the maximum values and the inter quartile range (Table 3.3). The difference between mean and median values illustrates a right skew in the data. The median value for read a depth of 42.58 is less influenced by the extreme values in the data set and is a more useful value to evaluate the centrality of the data set.

Table 3.3. Summary statistics reads for the average number of reads per target region per sample, in all populations and each experimental population (ALL = all populations; ART = artificial mixture population; BP = breeding population; ECO = ecotype population; IQR = Interquartile range; SD Standard deviation).

Population	Min	Median	Mean	Max	IQR	SD
ALL	0	42.58	180.03	7333.05	142.61	590.61
BP	0	39.53	179.03	7333.05	116.94	582.79
ECO	0	40.37	166.23	6338.49	112.01	578.78
ART	0	75.43	225.45	6325.21	142.61	682.99

In terms of representation failure within samples, some targets are missing among samples, as indicated by minimum values of 0 (Table 3.3). All targets were amplified within at least some of the samples, with no targets exhibiting maximum values of reads per locus of 0. The number of samples with a median depth of coverage of greater than or equal to the desired read depth of 50, is 333 samples, which is around 50.07% of the total samples. The number of targets with a median RD value across all samples is 87, accounting for 45.31% of total targets. No samples had a median RD higher than 500, while 95 had median value of 5 reads or fewer.

This experiment involved a number of groups, both in the target set (HD, CR, LG) and in the sample groups (BP, ECO, ART). It is useful to compare the distribution of reads between and within these groups to discern any biases within the target set or the sample groups. Between the marker sets, the LG had a median RD of 34.96, CR had a median RD of 26.34 and HD had the highest median read depth of 52.17. The 3 markers which were overrepresented all belonged to the HD marker set. All population groups exhibited similar minimum and maximum values (Table

3.3). There was a disparity observed among sample groups within median values however, with the highest values observed in the artificial mixtures population and comparable median values observed between the breeding population and the ecotypes.

3.3.2 Variants and missingness

Variant sets were generated for populations which were investigated for population structure, the breeding population, ecotypes and artificial mixtures. The filtering for genotyping quality, read depth and missingness sharply decreased the number of discovered SNPs (Table 3.4). The resulting SNPs were filtered to include target regions only. This also considerably reduced the number of SNPs, removing 76.53%-76.96% of variant sites. The number of target regions present in the final variant set ranged between 132-148 (68.75%- 77.08% of the original 192 targets). Variants were filtered to allow missingness of up to 50% to maximise capture of target regions. The overall missingness rate of SNPs for each population was lowest in the artificial mixtures, with the breeding and ecotype populations having similar rates of missing sites.

Table 3.4. Variant number at each stage of filtration for each experimental population, as well as the number of regions and SNPs per region in the final variant set and proportion of missing sites (BP = breeding population; ECO = ecotype population; ART = artificial mixture population).

	BP	**ECO**	**ART**
Raw SNPs	463,655	3,294,158	170,092
Filtered SNPs (quality parameters)	3,222	3,234	3,827
Filtered SNPs (target regions)	750	745	898
Target regions	134	132	148
SNPs per region	1-23	1-25	1-20
Proportion of missing sites	24.01%	24.28%	5.09%

3.3.3 Population delineation

Population structure was investigated in all three populations, the breeding population, the ecotype population and the artificial mixtures. The breeding population data set initially accounted for 473 individuals. Upon applying filters for variant sites, some individuals had 0 SNPs, so these individuals could not be

examined for further analysis. The number of individuals which were retained varied between populations (G1-G18), and in total 398 (n= 10-45) were retained. The filtered variant set for the remaining individuals included 750 SNPs from 134 target regions (Table 3.4). Each sample accounts for a single diploid individual with variants analysed as genotypes, utilising DAPC to detect population structure. The breeding population consisted of full sib families and cultivars. The clustering among full sib families demonstrates reasonably clear distinction between families (Figure 3.1). Individuals from the full sib families G11, G13 and G14 concentrated in clusters with one another, with the remaining 5 families clustering more independently. The families G13 and G14 shared a parental genotype (Portstewart X Fennema), and G13 and G15 also shared a parental genotype (Jumbo X Tyrone). This is further illustrated in Figure 3.2, where the membership probability of the assigned population for each sample was visualised. Among cultivars, individuals from each cultivar mostly converged with individuals from the same population (Figure 3.3). The most notable overlap between cultivar groups was observed between G10 and G1. The cultivars G1 and G10 exhibit lower proportions of correct assignment than the other cultivars while samples from G8, G9 and G4 were more definitively assigned to the correct group (Figure 3.4). All cultivars with the exception of G4 and G10 derive from separate breeding programmes.

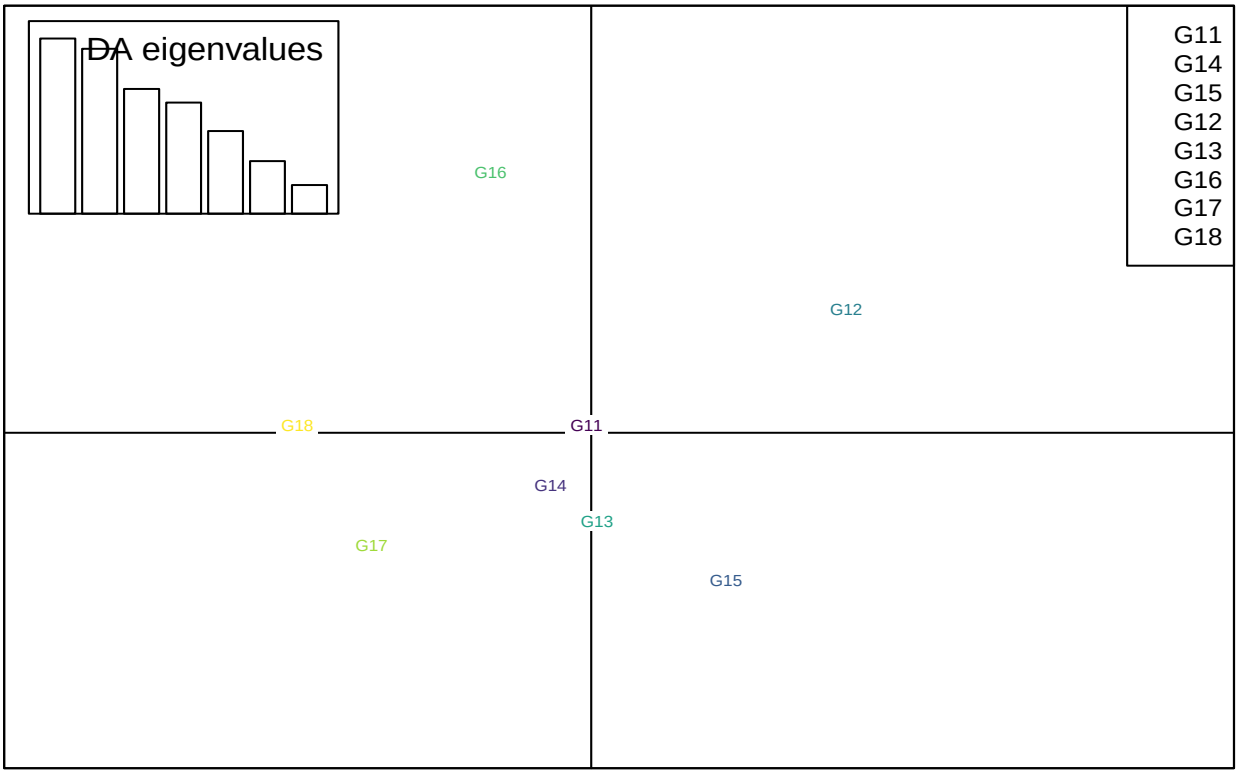


Figure 3.1. DAPC of full sib PRG families, utilising genotypes derived from 750 SNPs

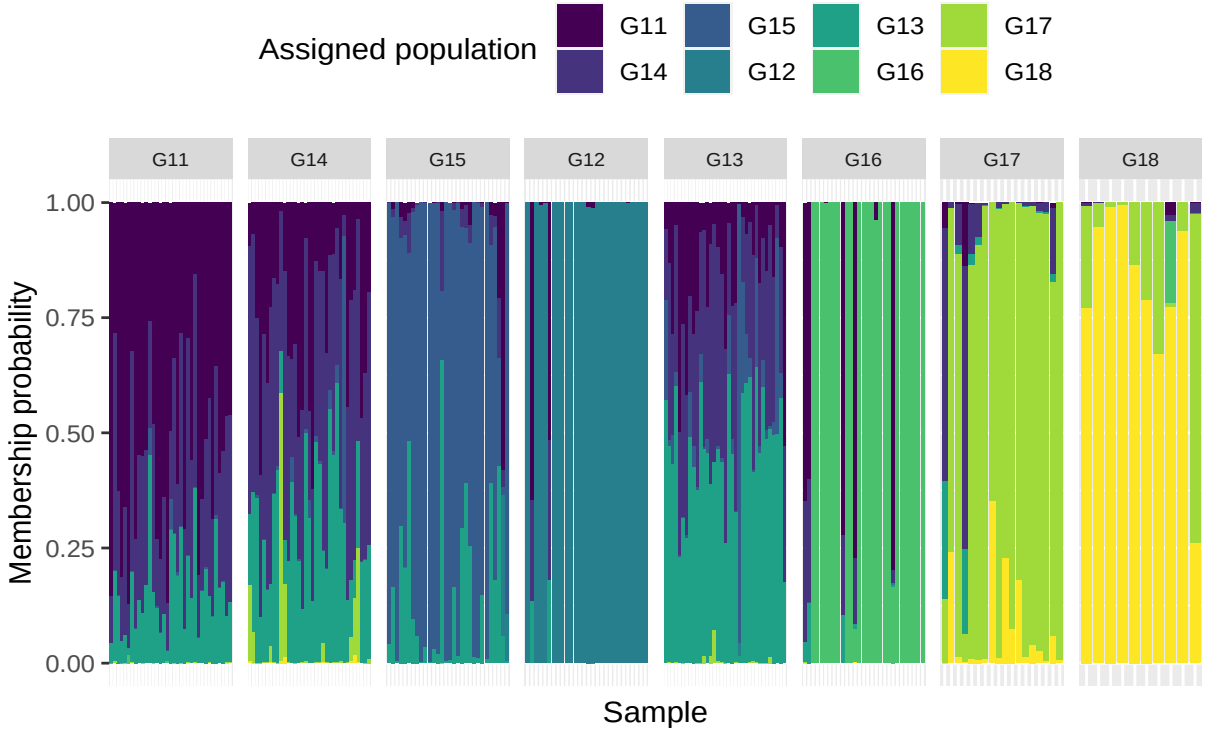


Figure 3.2. Membership probability of full sib PRG families, utilising genotypes derived from 750 SNPs

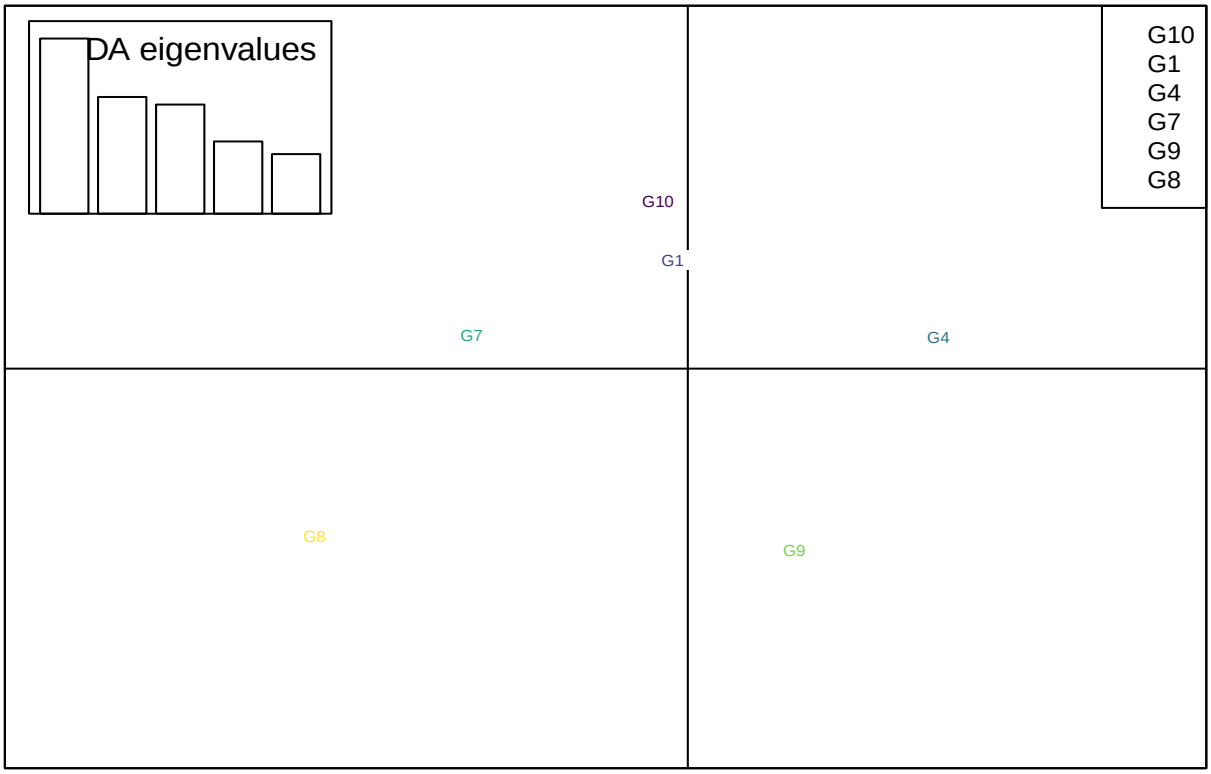


Figure 3.3. DAPC of PRG cultivars, utilising genotypes derived from 750 SNPs

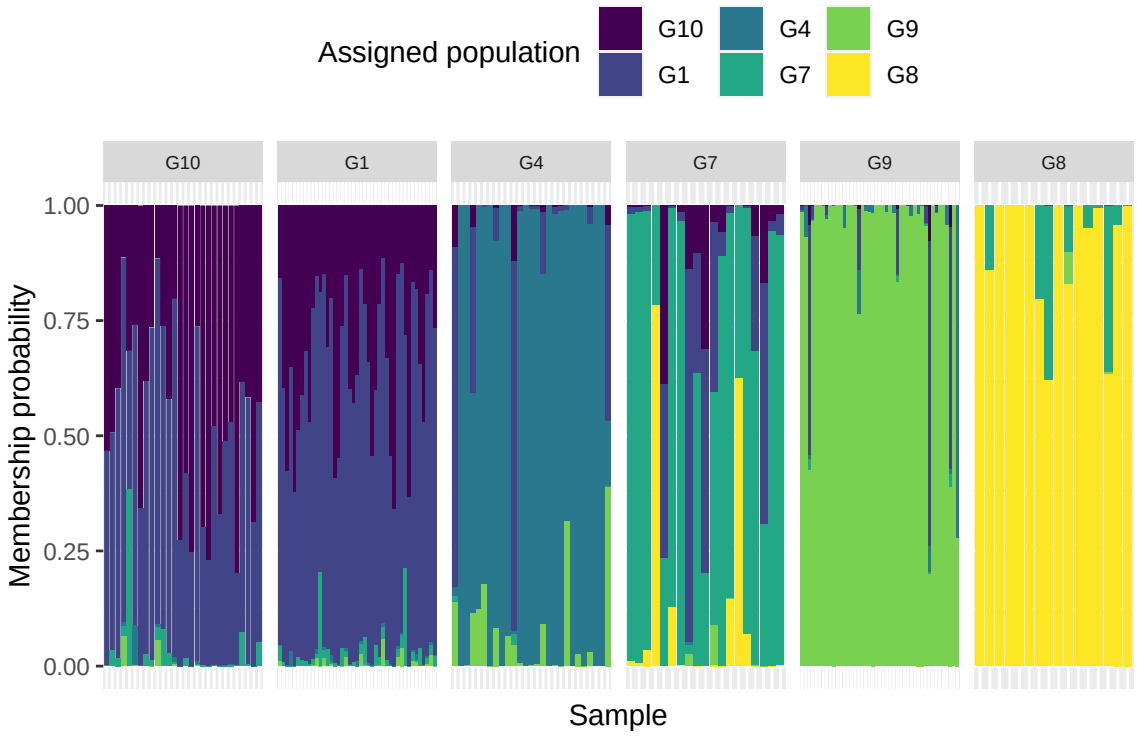


Figure 3.4. Membership probability of individuals from PRG cultivars, utilising genotypes derived from 745 SNPs

The ecotype population was investigated for the presence of discernible population structure utilising 745 SNPs present in 132 of the target loci (Figure 3.5). Each

sample accounts for a single diploid individual with variants analysed as genotypes, and as in the breeding population, DAPC was utilised for the analysis. Individuals of the cultivar Fennema were clustered separately from all other groups in this data set, and though individuals did not demonstrate a great likeness to one another, they were consistently and correctly assigned in terms of membership probability (Figure 3.6). Individuals from the second cultivar, Portstewart, did not display any differences from the majority of the ecotypes based on the DAPC. The ecotypes Wexford and Roscommon separated somewhat from the other ecotypes, however they were both observed to overlap with one another. The remaining 7 ecotypes clustered together without clear differentiation or structure. These trends are also observed in the population assignment, with the Wexford and Roscommon ecotypes sharing similar levels of assignment probability (Figure 3.6). The remaining ecotypes and the cultivar Portstewart display much greater levels of uncertainty population assignment.

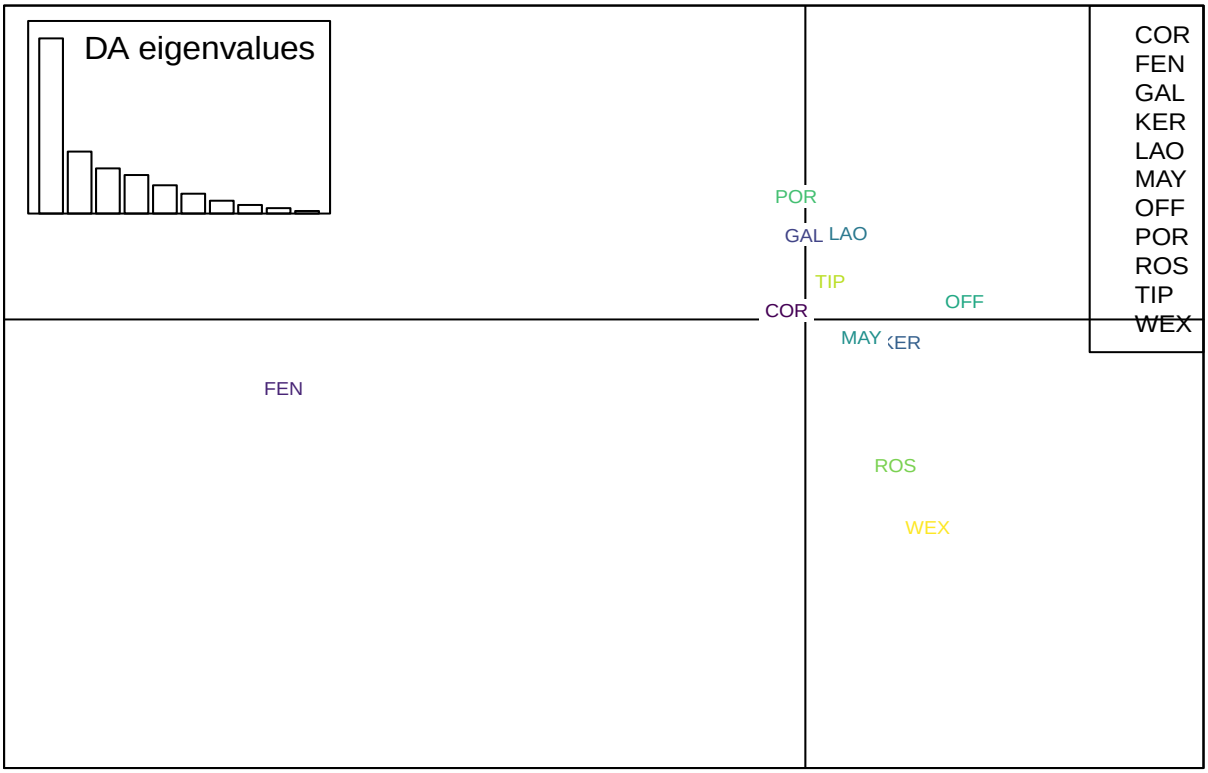


Figure 3.5. DAPC of ecotypes PRG populations, utilising genotypes derived from 745 SNPs

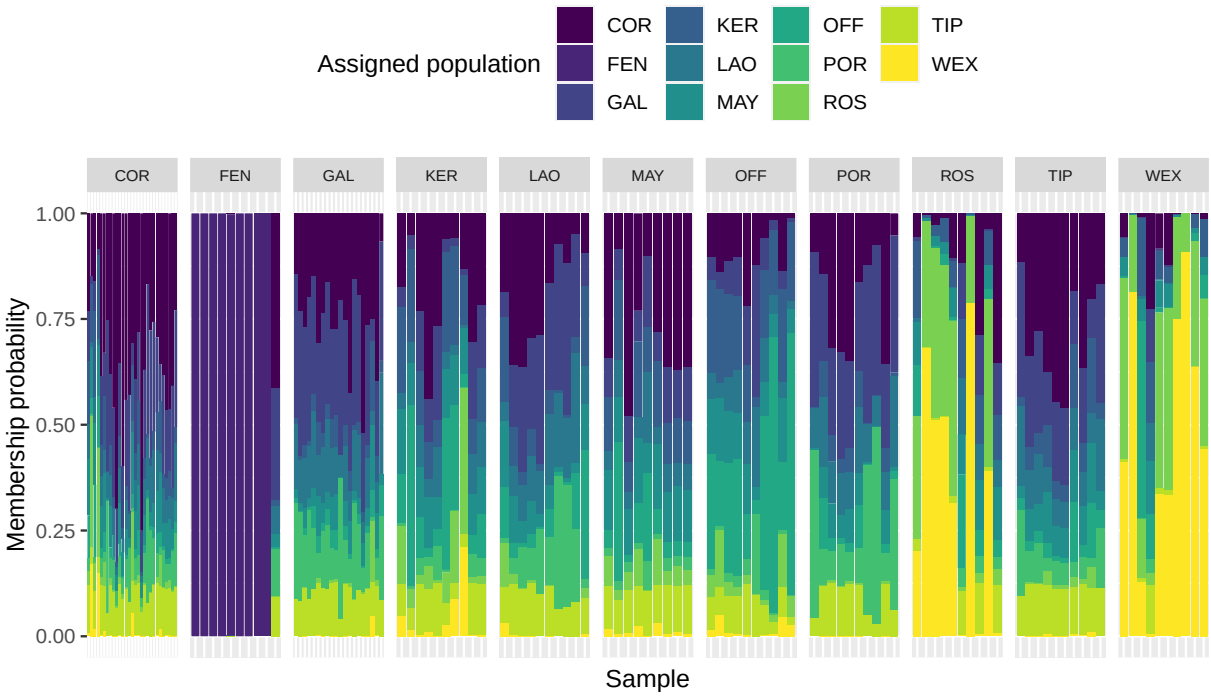


Figure 3.6. Membership probability of individuals from PRG ecotypes, utilising genotypes derived from 745 SNPs

The extent of population delineation in artificial mixtures was investigated, using 898 SNP frequencies located in 148 target regions, from 13 pure and mixed populations in varying ratios. PCA was performed to investigate the ability of the loci to detect population structure (Figure 3.7). Population structure was investigated with PCA as the DAPC package cannot handle population allele frequencies as samples. The first and second PCs represented 30.76% and 14.87% of the total variance respectively. The majority of the pure breeding families produced clear clusters, with replicates from families B, D, E, F forming independent groups. Some overlap between the pure populations was observed, namely between C and the 6 way mixture of uneven ratio, of which family C was the majority constituent. The uneven 6 way mixture also displayed some intersection with the 6 way mixture of even ratio.

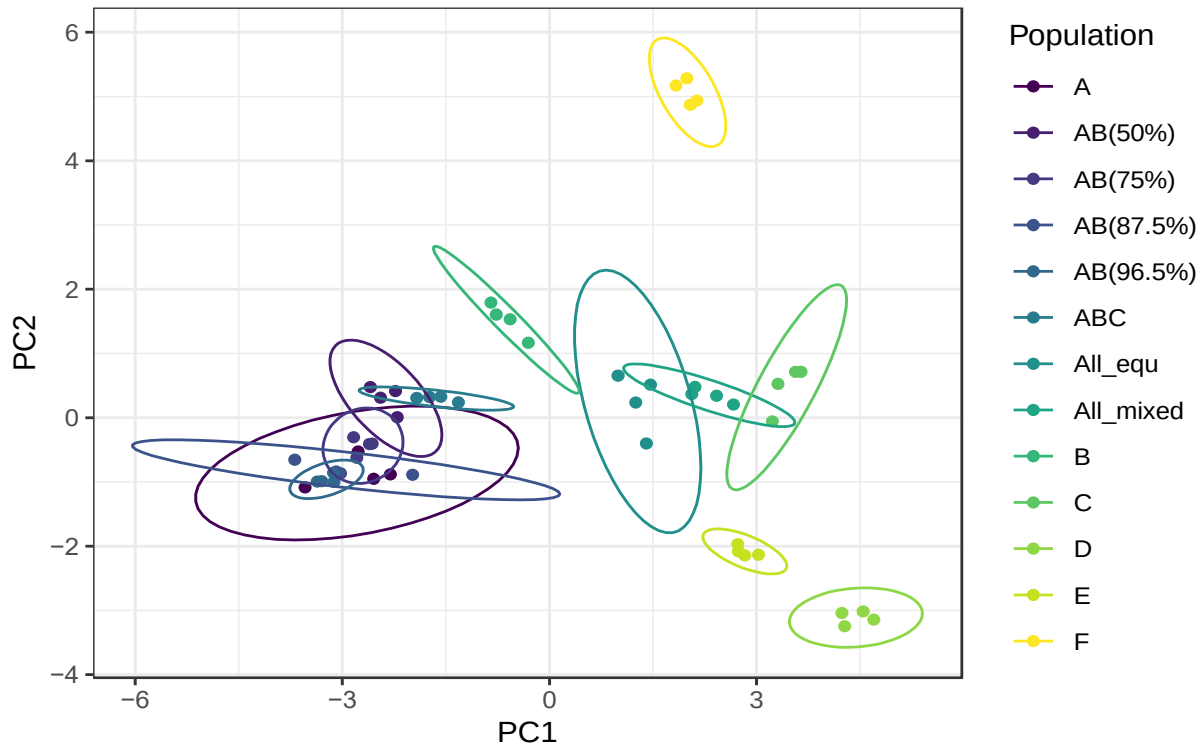


Figure 3.7. PCA of pooled pure families and artificial mixtures of F2 PRG populations, utilising allele frequencies of 898 SNPs

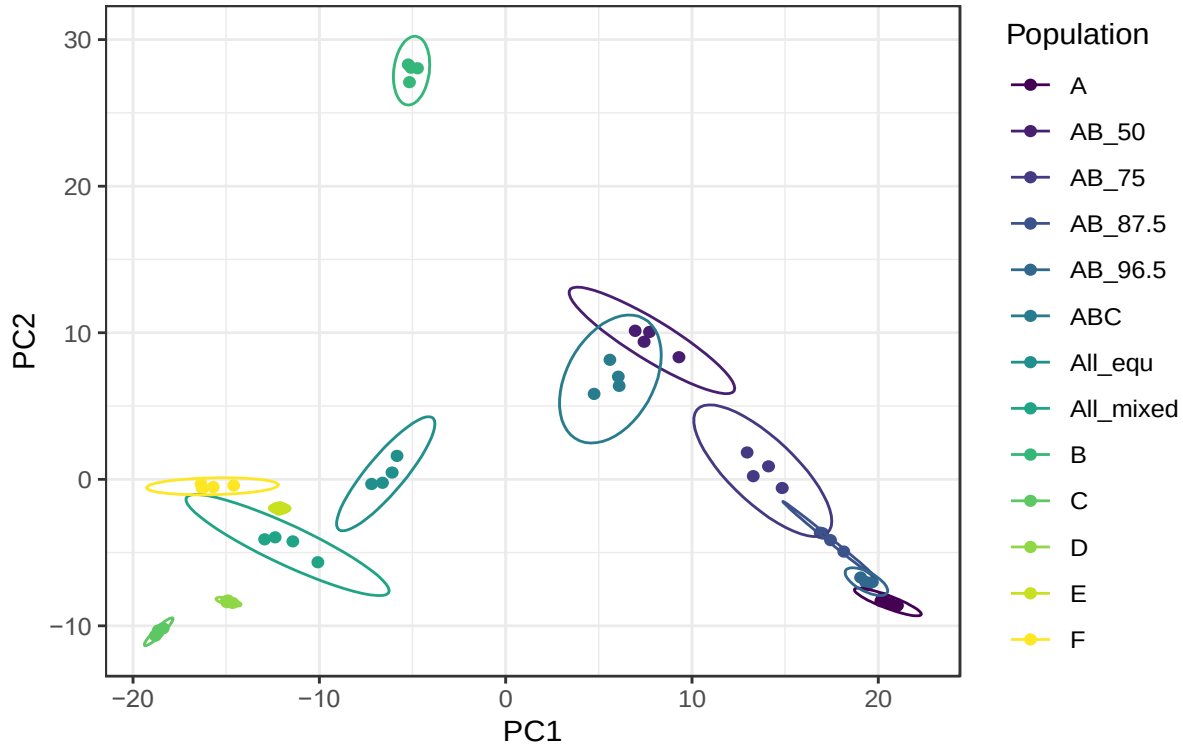


Figure 3.8. PCA of pooled pure families and artificial mixtures of F2 PRG populations utilising allele frequencies of 21,874 SNPs generated from GBS

The described population was also genotyped using a GBS approach, generating a putative SNP panel of 21,874 after filtering parameters were applied (bioinformatic analysis - up to the point of variant calling and filtering - performed by Dr Aude Perdereau, with subsequent analysis performed as part of this study). As in the GenoGrass approach, PCA was used to identify population structure with PC1 accounting for 36.38% of variance and PC2 accounting for 18.02%. This approach did yield greater delineation between the AB two way mixtures, however there was some assignment ambiguity in 3 of the pooled samples (Figure 3.8). Pure families that shared genotypes (D and E share the parent Jumbo X Tyrone, while E and F share the parent Portstewart X Fennema) are closer along with the six way combinations, mixtures of which they are components.

Each of the approaches was also analysed using hierarchal clustering to examine the relatedness (Figure 3.9 & 3.10). Similar to the trends observed in the PCA described above, two branches formed with pure families A and B and the two way and three way mixtures of which they are components and the remaing families grouping on a separate branch (Figure 3.9). This overall two branch structure was consistent with the structure observed in the GBS data set with the variations in lower branches (Figure 3.10)

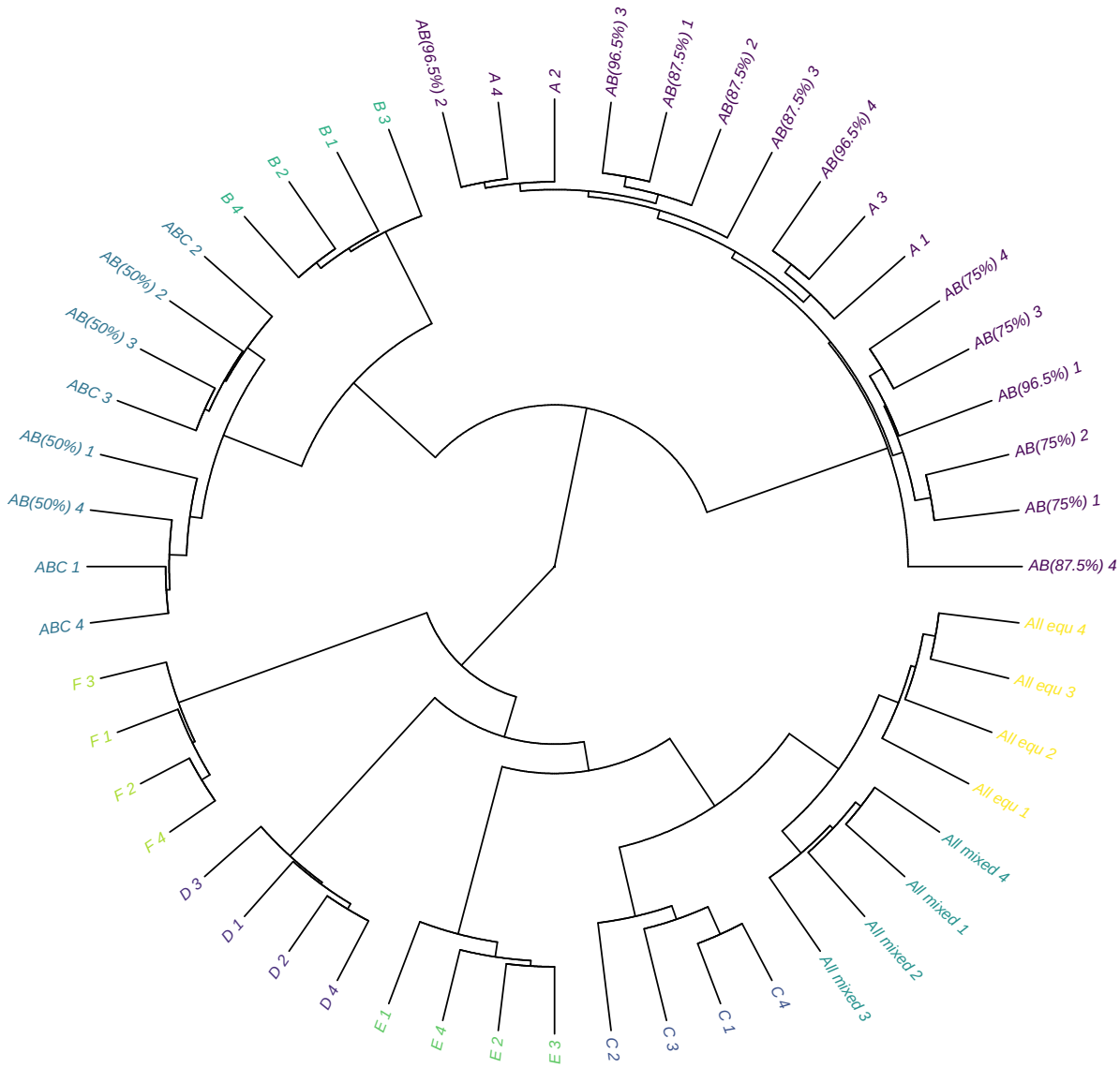


Figure 3.9. Hierarchical cluster analysis using Ward's minimum variance method of pooled pure families and artificial mixtures of F2 PRG populations, utilising allele frequencies of 898 SNPs generated from the reduced representation assay GenoGrass

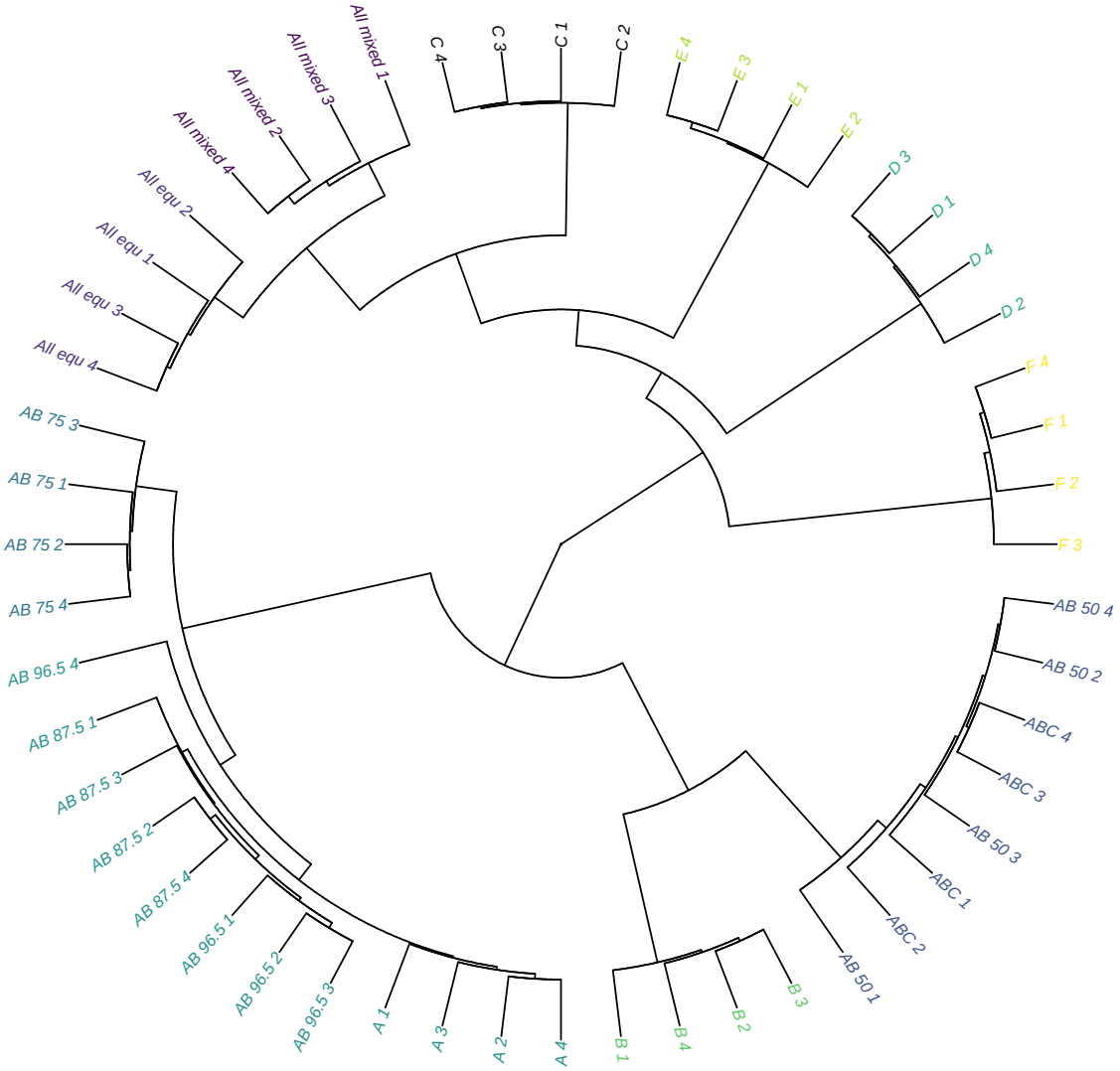


Figure 3.10. Hierarchical cluster analysis using Ward's minimum variance method of pooled pure families and artificial mixtures of F2 PRG populations, utilising allele frequencies of 21,874 SNPs generated from GBS

3.4 Discussion

Reduced representation genome sequencing is suitable for applications in crop breeding due to the relatively lower cost and the potential of fewer SNPs in breeding applications (Byrne *et al*., 2017; Zhang *et al*., 2020). We intended to demonstrate the utility of such SNP arrays in population delineation of number of population types to investigate the ability of the assay to detect proportional changes in PRG swards.

An adequate depth of coverage for target sequences is integral for the accurate ascertainment of genotypes and increasing RD in PRG pooled populations has been shown to improve reproducibility of allele frequencies (Byrne *et al*., 2013). Highly diverse species also benefit from enhanced coverage as a means of detecting rarer, lower frequency variants. Though the majority of targets were sequenced to a reasonable depth, some targets were absent, while some were clearly overrepresented. This should be taken into consideration in future target selection and primer design. It is likely that highly saturated targets influenced the overall evenness of the read distribution. The reduction in variant number of around 76% in all populations with the removal of SNPs despite prior quality filters, including mapping quality and depth of coverage, highlights the proportion of off target reads. Off target reads may have also reduced the overall efficiency of the assay. Amplification biases and artefacts can occur due disparate levels of GC richness, conserved vs polymorphic regions and repetitive sequences (Aird *et al*., 2011; Van Dijk *et al*., 2014). The locations of the targets included in the array have been identified from previous studies, which applied filters to extreme base compositions and repetitive regions. Though quality control measures were undertaken, issues may have arisen due to the completeness of the available genomic resources for PRG. These resources are crucial for the resequencing of the PRG genome, however the reference genome is currently in draft (Byrne *et al*., 2015). Some of the problems encountered can be mitigated with the continued improvement of genomic resources of PRG.

Performance of the assay was uneven across population types, with individuals from the breeding population and the ecotype exhibiting the least uniformity in coverage across individuals, while the artificial mixtures population exhibited higher median RD and less missingness. DNA from the artificial mixtures was extracted much more recently and though storage would have been adequate to preserve DNA quality unintentional thawing may have occurred through storage. DNA samples are archived and stored in communal freezers and inadvertent misuse can risk preservation. Quality checks were carried out to ensure a minimum levels of concentration and purity, though disparate levels between populations may have influenced amplification. Though it has been demonstrated that relatively few loci

can accurately detect minor changes in population composition, increased number of target regions may ensure an adequate number of sites are genotyped. Increased number of sites surveyed may also mitigate missingness in the sample set. In cucumber, so called “perfect SNPs” were designed where flanking regions of target sites are conserved to ensure consistent amplification and reduce missingness when using a reduced representation assay SNP-seq (Zhang *et al*., 2020). Approaches such as this can be undertaken to improve efficiency, however the lack of complete genomic resources in PRG hampers conclusive identification of such regions. Alternative reduced representation assays available such as SeqSNP (LGC Genomics, Berlin, Germany) and Ampliseq (Thermo Fisher Scientific, Waltham, MA, USA) have been used in crop research (Harper *et al*., 2019; Ogiso-Tanaka *et al*., 2019). SeqSNP utilises SPET to allow for an increased number of target regions while reducing the instance of primer interactions in target amplification (Barchi *et al*., 2019). This would allow for the scaling up of target number while and maintaining the cost and efficiency benefits of reduced representation assays.

Despite the associated technical challenges of the assay described above, this approach was successful in detecting population structure in several of the populations investigated. The delineation between full sib families and cultivars, along with other breeding material, has been previously described using GBS with many more times the number of SNPs (217,563) (Byrne *et al*., 2017). Individuals from full sib families were anticipated to be reliably delineated due to fewer parental genotypes resulting in less within group diversity. This was largely the case, with the reduced representation assay along with DAPC approach producing clear clusters with limited overlap between groups. The overlap that did occur was explained by the pedigree of the populations, with overlapping families possessing a common parent. Byrne *et al.* (2017) noted that cultivars derived from the same breeding programme (IBERS) exhibited clear separation from all other breeding material when individuals of varying diversity were genotyped using GBS and investigated for population structure using PCA. DAPC, with *a priori* group information, maximises variance the between clusters and reduces within group variance, producing clearer delineation than PCA. As cultivars are derived from a

greater number of parental genotypes, individuals from each cultivar were expected to more ambiguous groups than full sib families, though the results were largely comparable to that of the full sib families. Reducing the number of SNPs, along with employing an informed clustering approach can reliably delineate PRG full sib families and cultivars.

The ecotype population contained individuals from Irish ecotypes and two cultivars. Analysis was performed as in the breeding population. The reduced representation assay delineated some population structure and successfully delineated the cultivar Fennema population however all other groups clustered ambiguously. The populations from this group are derived from a study which investigated population diversity in PRG accessions and cultivars using SSR markers (Barth *et al*., 2017). Using SSR, the cultivar Fennema and the Laois ecotype were assigned to one of the two pools while the remaining populations assigned to the second pool. The Laois ecotype was not observed to separate from other ecotypes with Fennema when genotyped using SNPs, unlike the results of the original study. Wexford and Roscommon ecotypes showed some delineation, showing greater relatedness in terms of both genetic distance and population assignment than the other ecotypes investigated. This was also the case when the populations were genotyped using SSR markers (Barth *et al*., 2017). However they also exhibited a closer relationship with the cultivar Portstewart, which was not observed in the original study. Ecotypes are often a source of genetic diversity for breeding populations to incorporate novel traits into commercial cultivars, and are generally observed to possess high levels of heterozygosity within groups. However, Barth *et al*. (2017) found lower levels of heterozygosity, and speculated that this may be due to inbreeding, genetic bottle neck or directional selection. This indicates that within group diversity is unlikely to be influencing the inconsistencies between the two studies. Blackmore *et al.* (2015) investigated populations utilising a SNP array of 2185 variants and observed similar structure to that described in Barth et al. (2017). Though just 8 SSRs were utilised, the nature of multiallelic markers may be more advantageous to characterise population diversity among individuals. Multiallelic markers have been shown to be useful in investigating recombination events, ancestry and population structure

(Riday, 2011). The number of SNPs produced by this approach may not be adequate to characterise genetic diversity in ecotypes in comparison with SSRs.

The artificial mixtures population were characterised in this study both by the reduced representation assay GenoGrass and the with GBS. The complexity of the mixture types and the nature of the pooled samples was expected to increase difficulties associated with detecting population structure. Population structure was investigated with PCA as the DAPC package cannot handle population allele frequencies as samples. Utilising GenoGrass, the majority of pure families clustered clearly and groups which intersect contain common population constituents. Interestingly, the most complex, six way mixtures exhibited clearer association within groups than the two way mixtures. The two way mixture combinations, composed of families A and B, share two out of four parental types from the same breeding material. Likewise, in the three way A,B,C, combination, four of the six parental genotypes were derived from the same breeder, which may explain the ambiguity between this combination and that of the equal proportion two way mixture. Similar breeding material may be the source of ambiguity when proportions of a single component are increased using the method described. The results generated from the GBS for the artificial mixtures data set demonstrated greater success when differentiating the two way combination when just 25% of the minority constituent was included. The GenoGrass reduced representation utilises just 4.1% of the SNPs included in the GBS SNP set and in terms of pure mixture population delineation and complex 6 way mixtures, produced similar population structure. Though technical issues described above, including uneven coverage, and potential DNA degradation , reduced the number of usable target regions revision of and moderately scaled version of the assay would be suitable for further use in different aspect of perennial ryegrass breeding. Genotyping pooled perennial ryegrass populations was initially described by Byrne *et al.* (2013) where pooled populations of eight perennial ryegrass cultivars were successfully distinguished from one another using GBS. Verwimp *et al*. (2018) investigated the temporal changes in perennial ryegrass two way mixtures in the field using pooled samples, as well as validating the use pooled samples at a number of stages in the genotyping protocol. The number of unique individuals was known within the pool and therefore the

number of haploid genomes, informing analysis strategy. This is not always a convenient collection strategy particularly in the context of collecting field samples, however it is useful in certain aspects of analysis and may provide more complete assumptions about the resulting data set.

2.5 Conclusion

Reduced representation sequencing approaches lower the cost of sequencing per sample while maintaining the ability to differentiate genotypes generated. This will allow their use for a broad range of applications in PRG breeding. The characteristics of the PRG genome described can lead to technical challenges in sequencing experiments. We envisage a revised version of this assay being useful in PRG for identifying populations and the variants driving population delineation and subsequently providing markers for persistency studies.

Chapter 4

Changes in diversity of complex perennial ryegrass mixtures over time

Abstract

Swards of mixed cultivars of PRG are sown due to the perception of synergism between cultivars of complementary characteristics. This is despite the dearth of evidence indicating the benefits of mixed swards and studies demonstrating of the proportional instability of mixture components. Competition and selective pressures present in the field have the potential to affect the survival of individual plants in the sward and therefore alter proportions of mixture components. With the implementation of a low cost high throughput genotyping platform (SeqSNP), we investigated the changes in mixture components over time, reflective of mixtures sown on Irish farms. In two studies, two-way, three-way and four way mixtures were sown and monitored over a two year period. Seed from sown mixtures were retained and DNA from germinated seed was compared with DNA samples taken from the plants grown and grazed in plots over the subsequent two years. The first study investigated mixtures of contrasting ploidy, while the second study additionally investigated cultivar mixtures with contrasting maturity. SeqSNP was found to provide adequate depth of coverage and even representation across the selected target regions, allowing for the characterisation of genetic diversity over time. Both time point and seed combination were found to have a significant effect on genetic diversity (mean H_e). In Study 2 the original seed lots, were found to differ significantly from the other time points. Some shifts in genetic profiles were observed from the original seed, however these changes were sporadic and no clear trends or relationships were observed. The evidence from this study suggests that PRG mixtures were stable over two years post sowing.

4.1 Introduction

Perennial ryegrass (PRG) constitutes the majority of the grass sown in cultivated pastures in Ireland (Grogan & Gilliland, 2011). It is common practice for swards to be sown with three to four PRG cultivars (Culleton & Cullen, 1992; Gilliland *et al*., 2007). This was done to guarantee sward establishment, however, seed certification has mitigated sward failure and germination concerns are no longer a principal issue (Cowling & Lockyer, 1968). The practice has continued due to the perception and some evidence (Jones & Roberts, 1994) of mixed swards as acting synergistically under commercial grazing conditions. Tetraploid PRG varieties are often more attractive for forage due to their superior palatability and higher utilisation than diploid varieties (Castle & Watson, 1971; Tubritt *et al*., 2020). However, tetraploid cultivars have been observed to possess a lower tiller density and produce a more “open” sward (Neuteboom *et al*., 1988). As a consequence, tetraploid swards have been found to be susceptible to weed invasion and exhibit lower levels of persistency, here meaning the abundance stability of plants over time. The greater tiller density and regrowth capacity of diploid cultivars mitigates weed infestation and improves yield stability over time. Similarly, varieties with varying maturity are perceived to have beneficial effects on seasonal growth, due to the differences in growth rate between reproductive and vegetative tillers (Stapleton & Jones, 1987). Including such varieties in seed mixtures is believed to combine the positive attributes of each component. This is despite the conflicting evidence for synergism and the evaluation of PRG variety characteristics carried out in cultivar certifications and recommended lists as monocultures (Pollnac *et al*., 2014; Lowry *et al*., 2020; Gilliland & Gensollen, 2010; Tubritt *et al*., 2021).

Competition and selective pressures present in the field have the potential to affect the survival of individual plants in the sward and therefore alter proportions of mixture components (Clark, 2011). There is an indication that varieties with genotypes better suited to environmental pressures and management practices will exhibit enhanced persistence. Cashman *et al.* (2016) noted directional plant selection when applying frequent and infrequent cutting managements to twelve cultivars. Cultivars which possess a competitive advantage have been shown to influence proportional changes in binary mixtures (Griffith *et al*., 2016). This

indicates that as the proportions of sown cultivars change, their beneficial traits may not be represented in the same proportions over the life time of the sward. Morphological similarities between cultivars results in difficulties when discerning the contribution of each component within mixture (Quaite & Camlin, 1986). Monitoring changes in PRG genetic diversity over time using molecular markers can demonstrate the role of genetic diversity in sward longevity and persistency (Faville *et al*., 2020). In the case of binary mixtures, allozyme markers have been utilised previously, however this approach is limited in its ability to detect more subtle changes in complex mixtures (Griffith *et al*., 2016).

SNPs are abundant, scalable and adaptable markers and can be assayed using a range of sequencing approaches (Jones *et al*., 2009; Byrne *et al*., 2013; Arojju *et al*., 2018). SNPs have been used to characterise mixed, pooled PRG samples and have been demonstrated to track changes in sward diversity over time (Verwimp *et al*., 2018). Large scale sequence-based approaches often utilise enzymes to reduce genome complexity to capture SNPs and characterise pooled sample of PRG cultivars (Byrne *et al*. 2013). Targeted resequencing approaches such as GT-seq, Ampliseq and SNPseq (utilising Single Primer Extension Technology) allows genotyping of known SNPs with lower costs and greater precision than enzymatic approaches (Campbell *et al*., 2015; Ogiso-Tanaka *et a*l., 2019; Scaglione *et al*., 2019). SPET uses single primers - known as probes - for single direction amplification, though double probes can also be adapted to ensure the capture of novel sites as well as target sites (Scaglione *et al*., 2019). This approach benefits from *a priori* knowledge of the PRG genome (Byrne *et al*., 2015) to develop a low cost and robust method of characterising genetic diversity and identifying population structure in pooled PRG samples.

With the implementation of a low cost targeted genotyping approach, SeqSNP, we aimed to investigate the changes in mixture components over time, reflective of mixtures sown on Irish farms. We investigated the suitability of such an approach for characterising diversity over time. We did this by investigating a series of mixtures of varying complexities in two studies. Individual cultivar mixtures are difficult to test as a range of mixture combinations are commercially available. Consequently, it may be more useful to test mixtures in principle rather test each

combination independently. The first study included cultivars of varying ploidy with similar heading dates, to test the effects of ploidy alone. The second study involved cultivars of varying ploidy and differing heading date.

4.2 Methods

4.2.1 Plant material

Plant material used for analysis was obtained from seed samples of various cultivars and from field-grown trials derived from these seed sampled over two growing seasons.

Two field trials were sown side by side in August 2017, in Moorepark, Fermoy, Co. Cork. Both studies included plots measuring 3x5m^2 which were sown in a complete randomised block design, replicated by three (Figure 4.1). Each study consisted of four cultivars; Study 1 included cultivars with differing ploidy and similar heading dates and Study 2 with both varying ploidies and heading dates (Table 4.2). Seed combinations examined included two-way, three-way and four-way mixtures, including all cultivars, and resulting in 15 combinations per experiment (Table 4.2). Seed combinations were also investigated for agronomic performance, with trial management and fertilisation regime as part of a separate investigation (Tubritt *et al*., 2021). Each study underwent animal grazing and was grazed by dairy cows eight times from February to October in 2018 and 2019. High rates of fertilisation were applied annually: 310 kg N/ha, 30 kg P/ha and 40 kg K/ha. Higher rates were used to eliminate soil nutrition as a growth limiting factor.

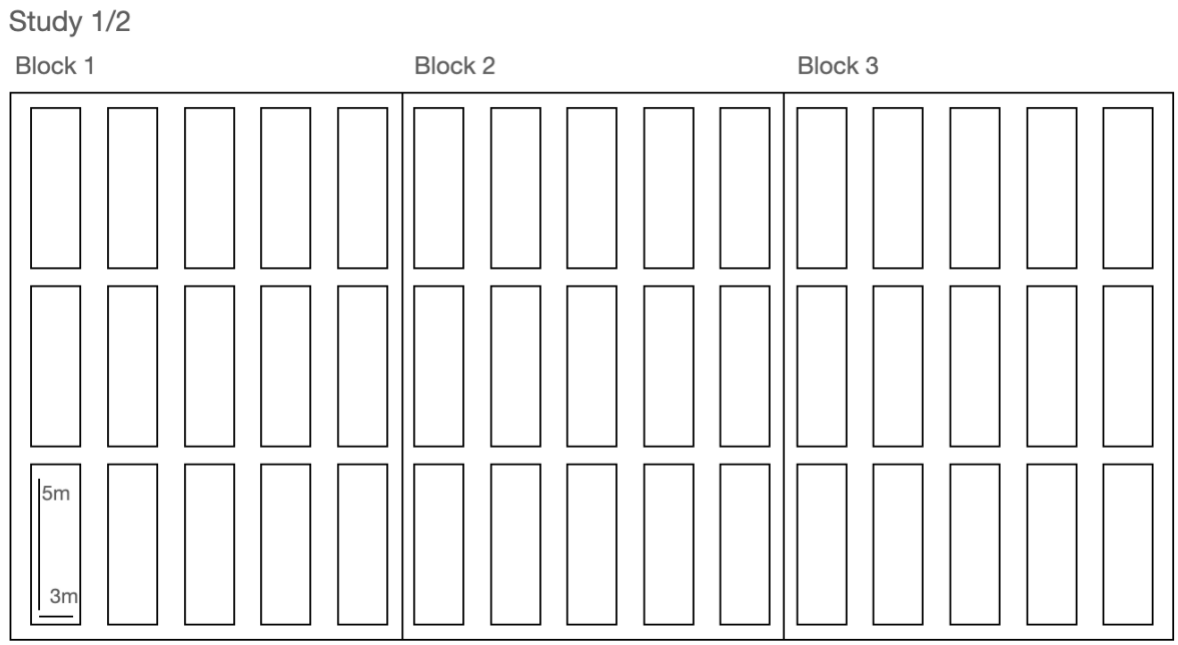


Figure 4.1. Study design

Sampling of the plots took place pre-grazing in 2018 as well as pre- and post-grazing in 2019 (Table 4.1). An autumn time point was excluded in 2018 due to an extended period of drought in the summer of 2018.

Table 4.1. Description of the time points used in this study

Time point	Code	Date
Original Seed	OS	Baseline
Spring 2018	S18	9th March 2018
Spring 2019	S19	13th February 2019
Autumn 2019	A19	29th October 2019

Plant material was collected from each plot using a 1m transect, with hand grabs of grass taken every 10cm. This resulted in 10 hand grabs per plot and collected hand grabs were bulked into single samples per plot and stored in -40ºC freezers. Hand grabs are used to simulate the herbivory of a bovine animal. This allows the sampling of the most productive plants and the therefore represents plants with the greatest biomass rather than abundance. The total plant material per plot was ground in pestle and mortar with liquid nitrogen to a fine powder and 100µl of ground material was deposited into 2ml microtubes. DNA was extracted using the CTAB extraction method (Doyle, 1991). For the original seed lots, 0.5 of seed from two of

the three plot replicates was subsampled from the seed mixtures sown (OS time point). Representative samples were ensured by homogenising seed mixtures pre-sampling. Seed was germinated in 9x9cm pots and plant material was collected as described in Byrne *et al.* (2013). DNA was isolated, using the method described above and samples were quantified and normalised for sequencing.

Table 4.2. Description of the cultivars and mixture components sown in Study 1 and Study 2. Ploidy (D = Diploid, T = Tetraploid) and maturity (I = Intermediate, L = Late). Each mixture contains an equal proportion of each seed component. Cultivar superscripts denote the following breeders; 1= IBERS; 2 = Teagasc; 3 = DSV; 4 = AFBI.

Study 1			**Study 2**		
Cultivar	**Ploidy**	**Maturity**	**Cultivar**	**Ploidy**	**Maturity**
Pure			*Pure*		
Aberchoice[1]	D	L	Aberchoice[1]	D	L
Oakpark[2]	D	L	Aberclyde[1]	T	I
Astonenergy[3]	T	L	Astonenergy[3]	T	L
Solas[2]	T	L	Rosetta[4]	D	I
Two way			*Two way*		
AbC/Oak	D:D	L	AbC/Ros	D:D	L:I
AbC /Aston	D:D	L	AbC/Aston	D:T	L:L
AbC /Sol	D:T	L	AbC/Aberclyde	D:T	L:I
Aston/Sol	T:T	L	Aberclyde/Aston	T:T	L:I
Aston/Oak	T:D	L	Aberclyde/Ros	D:T	I:I
Oak/Sol	D:T	L	Aston/Ros	D:T	L:I
Three way			*Three way*		
AbC/Oak/Sol	D:D:T	L	AbC/Aston/Ros	D:D:T	L:I:L
AbC/Oak/Aston	D:D:T	L	AbC/Aberclyde/Ros	D:D:T	L:I:I
AbC/Aston/Solas	D:T:T	L	AbC/Aberclyde/Aston	D:T:T	L:I:L
Oak/Aston/Solas	D:T:T	L	Aberclyde/Aston/Ros	D:T:T	L:L:I
Four way			*Four way*		
AbC/Oak/Aston/Sol	D:D:T:T	L	Abc/Aberclyde/Aston/Ros	D:D:T T	L:I:L:I

Study 1 cultivar abbreviations: Aberchoice = AbC; Oakpark = Oak; Astonenergy = Aston; Solas = Sol
Study 2 cultivar abbreviations: Aberchoice = AbC; Aberclyde = Aberclyde; Astonenergy = Aston; Rosetta = Ros

4.2.2 Genotyping protocol

Sequencing was carried out using a reduced representation assay, SeqSNP (LGC, Berlin). SeqSNP utilises SPET to capture target regions, and double probes were employed to mitigate missing data. A total of 500 targets were selected for their even distribution throughout the PRG genome, with 495 using double probes and the remaining five targets using single sequence probes. Single targets were used to capture a small number of targets due to multimapping of second probes. Target regions which had also been successfully used in the previous experiment (Chapter 3) were included in the target panel. Target regions were 40-153bp (mean = 105.33) in length.

4.2.3 Data preparation

Demultiplexing was carried out using the Illumina bcl2fastq v2.20, allowing for 1 to 2 mismatches. Reads were clipped of adaptor sequences and trimmed for poor quality bases and length (carried out as part of sequencing service by LGC, the remaining steps were carried out by NC). Paired reads were aligned to the PRG reference genome using bwa v 0.7.17 (Li & Durbin, 2010).. Depth of coverage for each target site was determined using the SAMtools v. 1.2 (Li *et al.*, 2009), coverage function, where a bed file containing target positions was used to generate the number of reads per site. Variants were called using vcftools v. 0.1.16 mpileup (Danecek *et al*., 2011). Variant sites with a genotype quality of at least 30, a read depth of more than 10 and missingness of less than 50% were retained. The read depth and missingness thresholds were reasonably relaxed to capture as many target sites as possible.

4.2.4 Data analysis

Allele frequencies for variant sites were calculated in R v. 4.0.1 (R Core Team, 2021). Principal Component Analysis (PCA) was used to investigate overall data structure, that was structured by cultivar mixture and locus, with allele frequencies at each. Changes in trends over time, was investigated using the *prcomp* function in R; it is a core function in R using singular value decomposition which examines the covariances / correlations between individuals. The expected heterozygosity (H_e) was calculated for each variant site ($H_e = 2 \times q \times (1 - q)$) using R and

averaged for each plot and the four time points. The following model was used to investigate the influence of seed combination and year on the mean H_e per plot sample:

$$H_{ijk} = \mu + S_i + T_j + R_k + (S_i \times T_j) + \epsilon_{ijk}$$

Where H_{ijk} denotes the mean H_e for the seed combination (S) i (1-15), for time point (T) j (1-4), replicate (R) within a block k (1-3), where μ is the mean and ϵ_{ijk} denotes the residual error. Post-hoc testing for significant differences between means was carried out using the *emmeans* package (Lenth, 2020) in R. Assumptions for linearity, homoscedasticity and independence analysis were checked and met.

4.3 Results

4.3.1 Read alignment and missingness

Sequencing produced an average read depth of 200x. Filtering for genotyping quality, read depth and missingness sharply decreased the number of discovered SNPs (Table 4.3). The resulting SNPs were filtered to include target regions only. This marginally reduced the number of SNPs, removing 7.34-8.08% of variant sites. The number of target regions present in the final variant set ranged between 426-429 (85.2%-85.8% of the original 500 targets). Variants were filtered to allow missingness of up to 50% to maximise capture of target regions.

Table 4.3. Variant number at each stage of filtration for each experimental population, as well as the number of regions and SNPs per region in the final variant set and proportion of missing sites

	Study 1	**Study 2**
Raw putative variants	39,947	38,472
Quality filtered variants	1,026	1,036
Location filtered variants	943	960
Target regions	429	426
Missing	12%	18%

The overall missingness rate of SNPs for each population study was 12-18%. A disproportionate level of missingness occurred in a number of samples. Samples with fewer than 50% of polymorphic sites were removed due to the inaccuracies associated with analysing samples with high rates of missingness (O'Leary *et al*., 2018). The number of samples removed from Study 1 was 15 and 28 were removed

from Study 2. Once these samples were removed, the missing rate reduced to 4.51% and 5.19% in Study 1 and Study 2 respectively.

4.3.2 Differentiation among populations based on PCA

PCA was utilised to detect the most obvious trends in delineation. PC1 accounted for 23.6%, while PC2 accounted for 16.3% of the variance observed in Study 1 (Figure 4.2 A).

In Study 1, the overall trends indicated greater representation of one to two cultivars in seed mixtures in the original seed OS time point. This was then followed by movement to an intermediate distance between constituent cultivars at later timepoints. This is illustrated in the Study 1 two-way combinations of cultivars with contrasting ploidies, where mixtures appear to cluster more closely to their tetraploid component (Aston) in the OS time point and progressively move to an intermediate distance between cultivars in later time points (Figure 4.2 B). This was with exception of the Oak/Sol combination, where mixed populations consistently clustered closely together with both Oakpark and Solas cultivars, which are derived from the same breeding programme (Teagasc) and were also observed to cluster closely with each other (Supplementary Figure S4.2). In the pure tetraploid combination, Aston/Solas, Solas dominated the earlier population while in the pure diploid, AbC/Oak, mixed population replicates displayed variation within the population with greater likeness to Aberchoice at later timepoints (Appendix 11).

Three-way mixtures, unlike the two- and four-way mixtures, include combinations of overall unequal proportions of ploidy. Similar trends were observed in three-way mixes as in the two-way mixtures; greater distances within populations were observed in the original seed batches in the OS timepoint, with clearer clusters emerging in later timepoints. In both the diploid and tetraploid dominated mixtures, AbC/Aston/Solas and AbC/Oak/Aston respectively, mixtures at the OS time point were more alike to the tetraploid component and moved towards an intermediate distance between all components in later timepoints (Figure 4.2 C; Appendix 12). In the majority diploid mixture AbC/Oak/Solas, mixed seed batches in the OS timepoint clustered more closely to the both the Teagasc bred cultivars, Oakpark and Solas, and drifted towards an intermediate distance between all constituent cultivars (Appendix 12). In the second majority tetraploid mixture, mixed populations were

observed to cluster at an intermediate distance from all component clusters with some mixed population samples drifting more closely to the Teagasc cultivars, Solas and Oakpark (Appendix 12).

The single four-way mixture was relatively stable over the four timepoints. It clustered more closely with the Teagasc bred (Oakpark, Carlow) cultivars, though this may also be observed as an intermediate distance between the Astonenergy and Aberchoice (Figure 4.2 D).

In Study 2, PC1 accounted for 21.1%, while PC2 accounted for 19.1% of the observed variance (Figure 4.3 A). In Study 2, two-way mixtures displayed similar trends to that in Study 1, with mixed populations often clustering more closely to one cultivar, then shifting to a moderate distance between cultivars in later timepoints. In two of the mixtures with contrasting ploidy, AbC/Aberclyde and AbC/Aston it was often the tetraploid component that was dominant at the OS timepoint, drifting to an intermediate distance between component cultivars (Appendix 13). The other two-way mixtures with contrasting ploidy, Aberclyde/Ros and Aston/Ros, were relatively stable across time points, though demonstrated a greater likeness to the Rosetta cultivar (Appendix 13; Figure 4.3). The tetraploid only mix, Aberclyde/Aston, consistently displayed moderate likeness between cultivars throughout all timepoints, despite contrasting maturity (Appendix 13). The late heading cultivar Aberchoice in the diploid only mix, AbC/Ros, displayed greater dominance in the original seed batches and then moved to an intermediate distance between cultivars post sowing (Appendix 13).

As in Study 1, three-way mixtures include overall combinations of unequal proportions ploidy. Study 2 has the added variable of maturity and as with ploidy mixtures are comprised of unequal proportions of each of these variables (Table 4.2). The majority tetraploid, late heading mixture AbC/Aberclyde/Aston displayed greater similarities to the tetraploid late heading cultivar Aston (Appendix 14). Regardless of ploidy/maturity the mixtures AbC/Aberclyde/Ros, AbC/Aston/Ros, Aberclyde/Aston/Ros displayed greater similarities the diploid intermediate cultivar Rosetta, then shifted to display equal affinity to other mixture components (Figure 4.3 C; Appendix 14).

As in Study 1, the single four-way mixture in Study 2 exhibited relative stability over the four timepoints. The four-way mixture clustered more closely with the late heading cultivars Rosetta and Astonenergy and shifted towards Aberclyde in later time points (Figure 4.3 D).

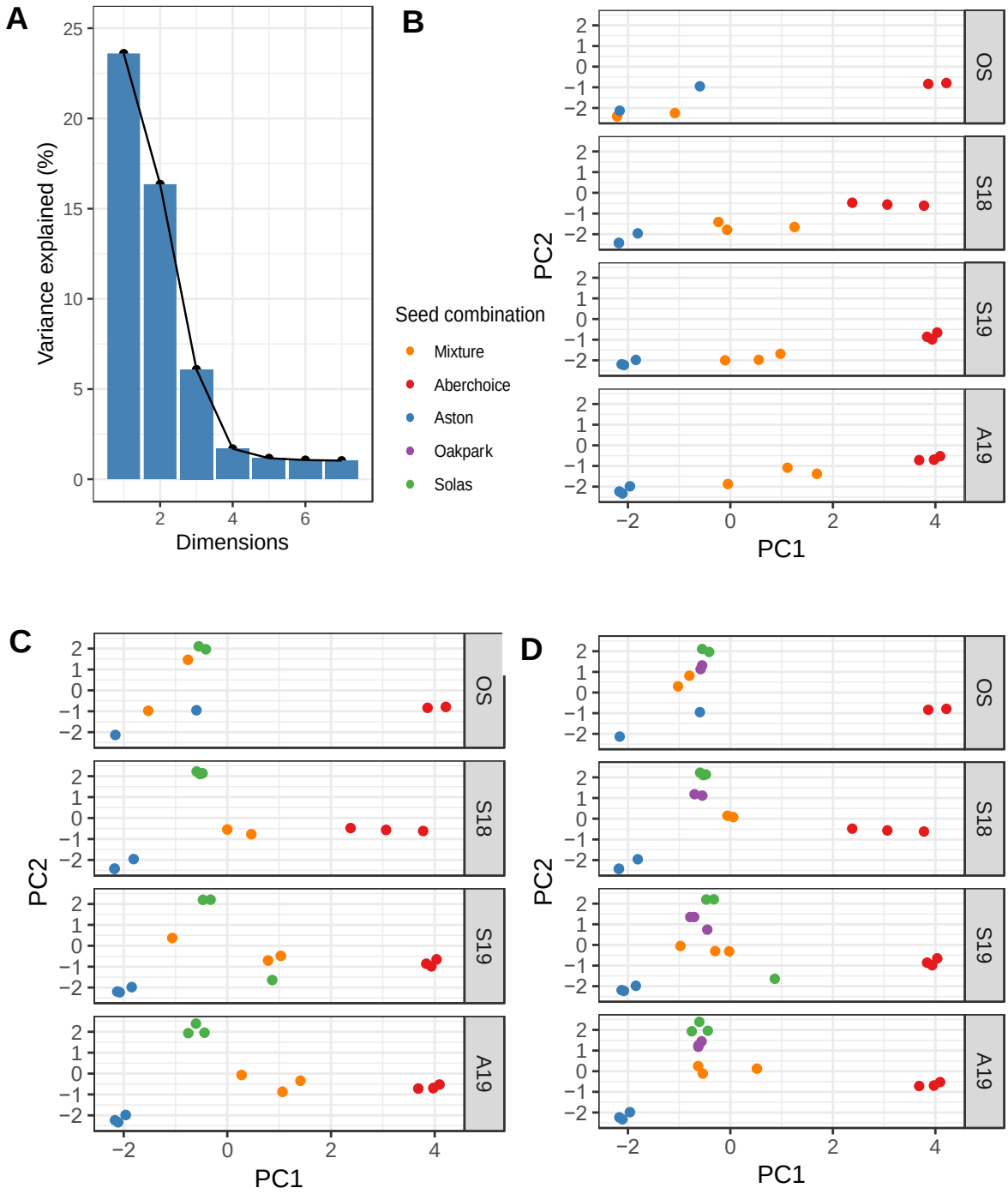


Figure 4.2. Percentage of variance explained by each principal component in Study 1 (A), PCA displaying first two principal components of a two-way mixture (B), three-way mixture (C) and four-way mixture (D). Mixtures consist of cultivars of varying ploidy: Aberchoice (diploid, Oakpark (diploid), Astonenergy (tetraploid), Solas (tetraploid).

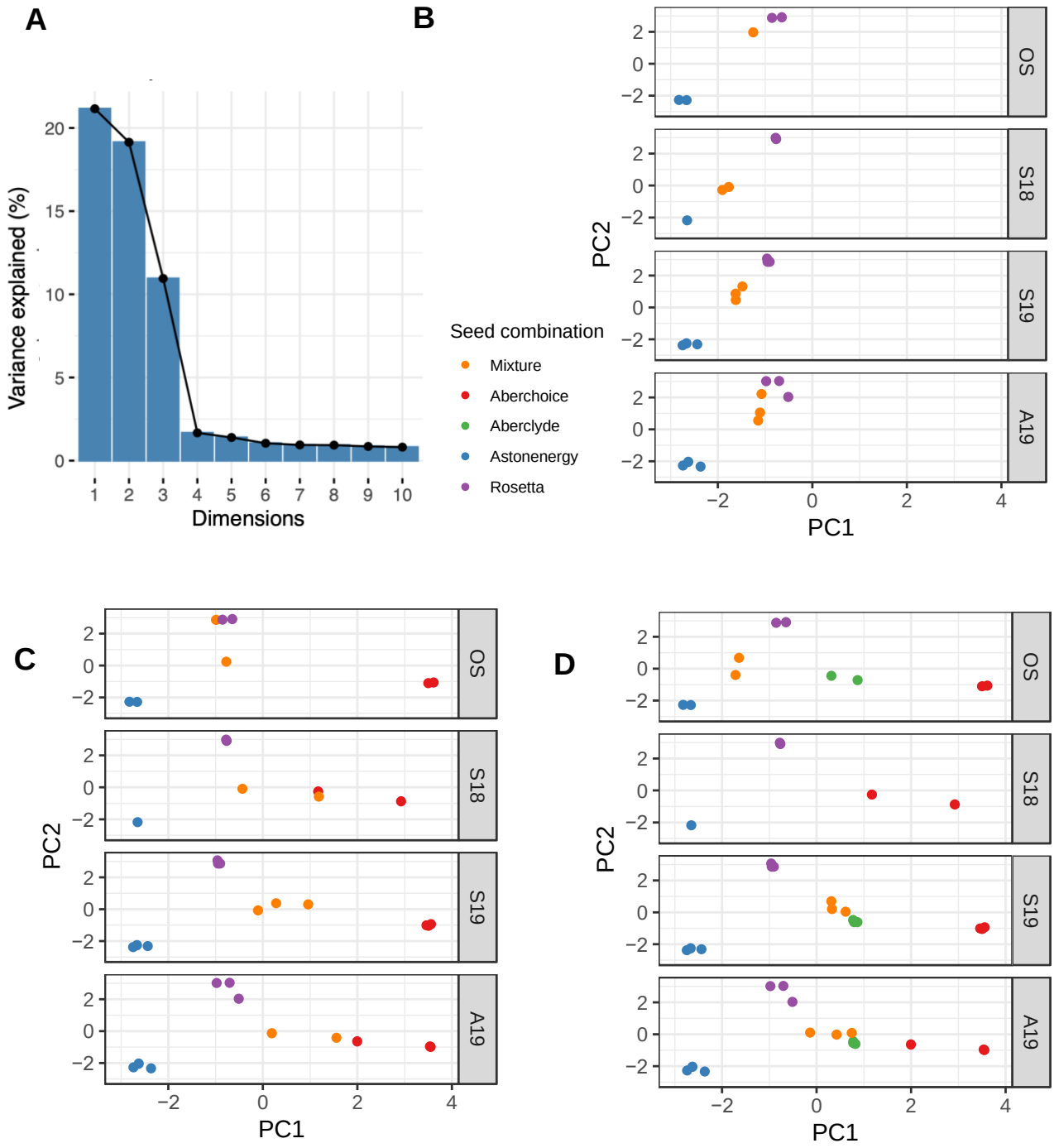


Figure 4.3. Percentage of variance explained by each principal component in Study 2 (A), PCA displaying first two principal components of a two-way mixture (B), three-way mixture

(C) and four-way mixture (D). Mixtures consist of cultivars of varying ploidy and maturity: Aberchoice (diploid, late maturity) Aberclyde (tetraploid, intermediate maturity), Astonenergy (tetraploid, late maturity), Rosetta, (diploid, intermediate).

4.3.3 Changes in genetic diversity over time

Changes in heterozygosity can indicate selection as reductions in levels of mean H_e suggests reduction in genetic diversity (Verwimp *et al*., 2018). The effect of seed combination and time were investigated to discern the effect of these variables on expected heterozygosity and the possible interaction they may have with one another. Seed combination accounts for the greatest proportion of variance in mean H_e. Time point had a significant effect on H_e; also the interaction between timepoint and population (seed combination) was significant (Table 4.4). Post-hoc analysis was conducted to discern which time points differed significantly from one another, however no timepoints differed significantly (all p>0.05). The significance level of the comparisons was corrected for multiple testing.

Table 4.4. ANOVA results of the mean H_e based on 943 SNPs for Study 1

Variable	**d.f**	**SS**	**MSS**	**F**	**P**	**Sig.**
Replicate	2	0.000	0.000	0.340	0.775	ns
Time point	3	0.000	0.000	2.834	0.042	*
Seed combination	14	0.002	0.000	27.642	<2e-16	***
Time point:Seed combination	42	0.002	0.000	1.652	0.025	*
Residuals	89	0.002	0.000			

***(p< 0.001); **(p< 0.01); *(p< 0.05); ns = not significant

In Study 1, the mixed seed combinations displayed the highest levels of heterozygosity, with some exceptions in the original seed batches. Oakpark and Solas had the highest levels of heterozygosity of the pure seed batches (Figure 4.4 C). Many of the mixed seed combinations increased in mean H_e from OS to S18, followed by declines in heterozygosity subsequently (Figure 4.4; Appendix 15). Though not all mixtures followed these trends and though mean H_e fluctuated over time. This was particularly the case for pure cultivars (Figure 4.4; Appendix 16).

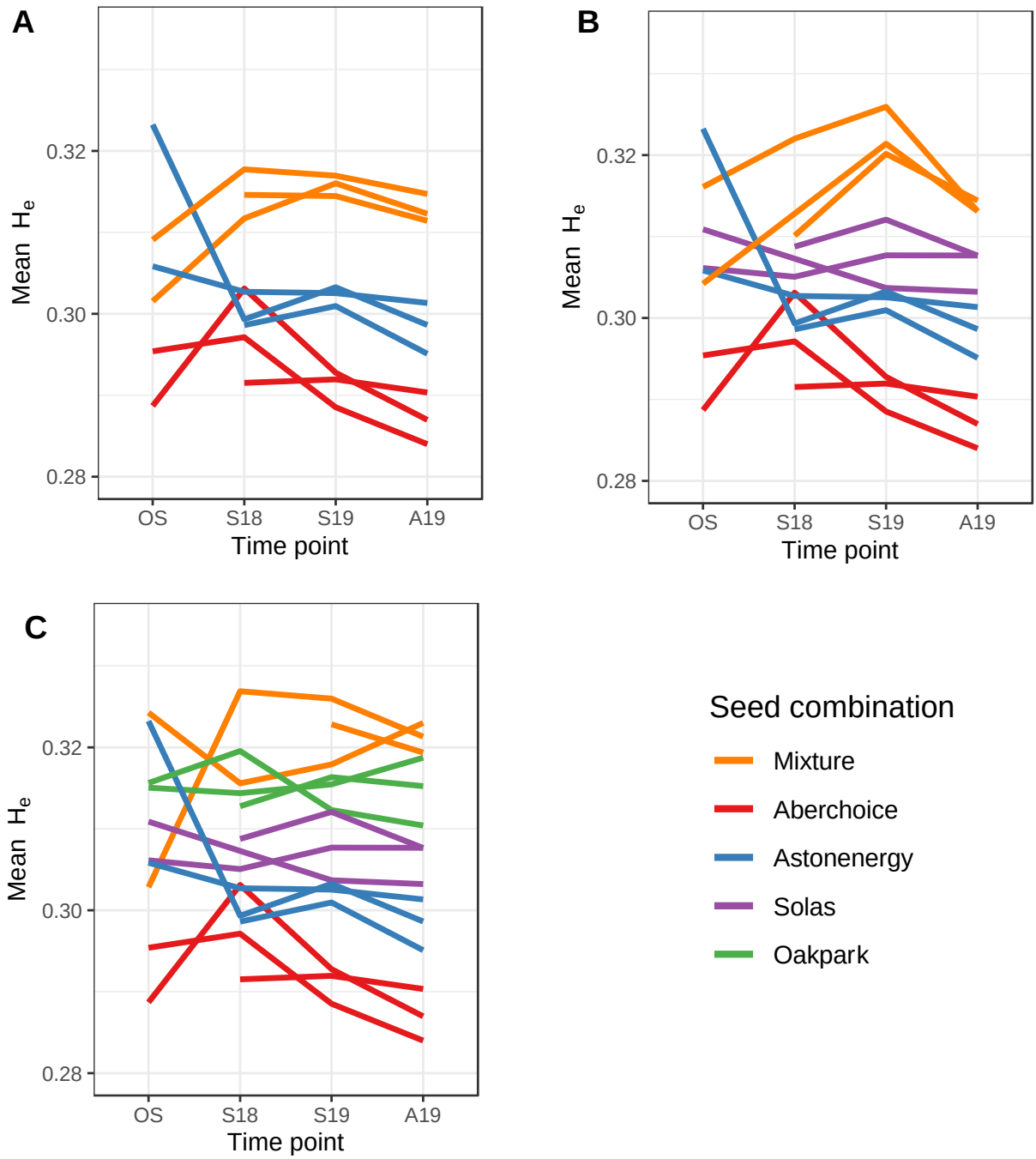


Figure 4.4. Study 1: Trends in mean H_e over time in two-way mixtures (A) three-way mixtures (B) four-way mixtures (C). Mixtures consist of cultivars of varying ploidy: Aberchoice (diploid, Oakpark (diploid), Astonenergy (tetraploid), Solas (tetraploid).

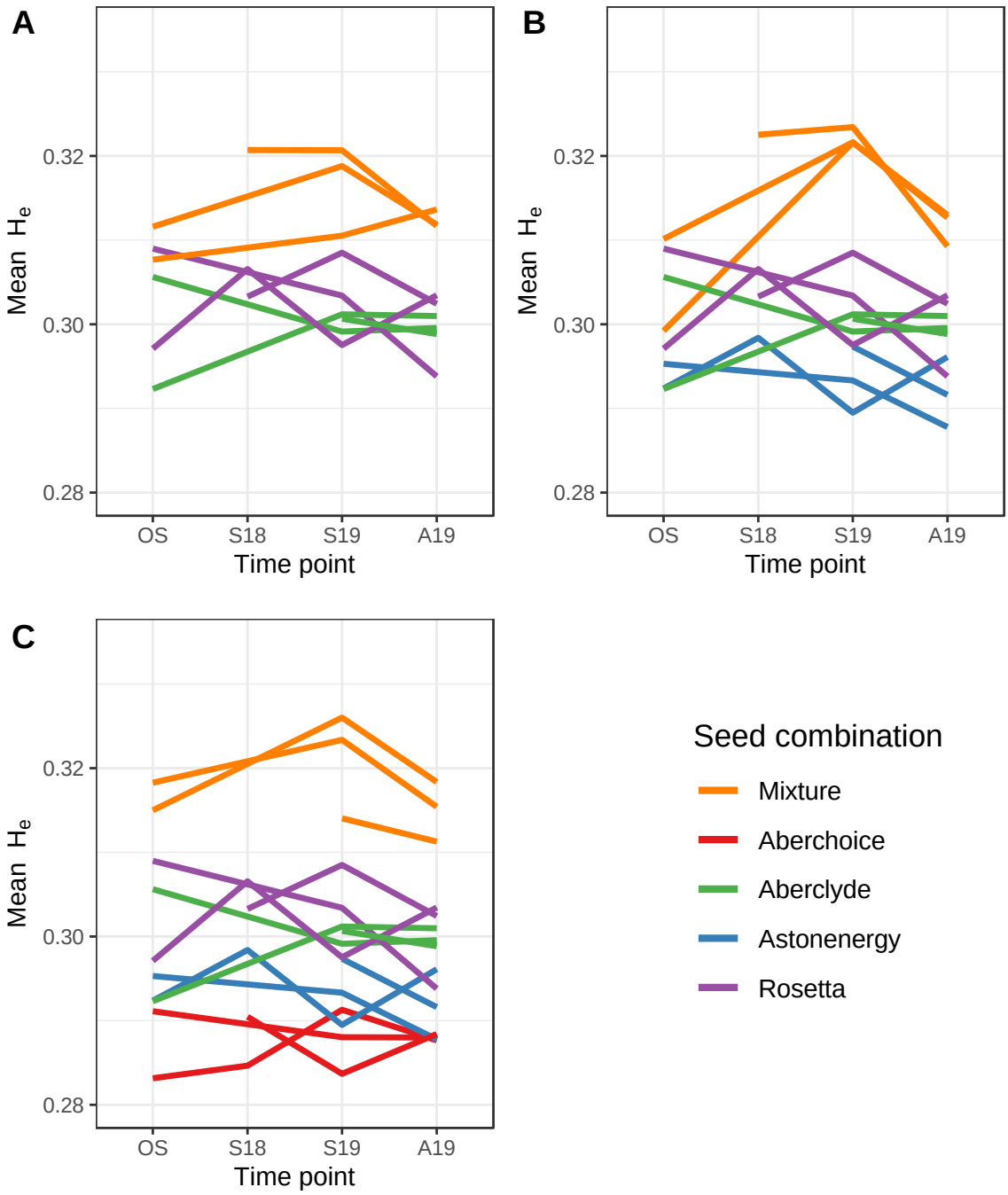


Figure 4.5. Study 2: Trends in mean H_e over time in two-way mixtures (A), three-way mixtures (B) and four-way mixtures (C). Mixtures consist of cultivars of varying ploidy and maturity: Aberchoice (diploid, late maturity) Aberclyde (tetraploid, intermediate maturity), Astonenergy (tetraploid, late maturity), Rosetta, (diploid, intermediate).

A similar model was applied to Study 2. As in Study 1, seed combination accounts for the greatest proportion of variance in mean H_e and was found to have a significant effect on mean H_e (Table 4.5). Time point was also demonstrated to affect mean H_e significantly. Upon inspection of post hoc analysis results, OS varied significantly from all other time points and was the only time point to do so ($p<0.05$). S18 was excluded in the comparisons due to missing data. Unlike in the Study 1, the interaction between time point and seed combination in Study 2 did not significantly influence the variation in mean H_e (Table 4.5).

In Study 2, the mixtures demonstrated the highest levels of heterozygosity, with some exceptions, largely in the original seed batches. Rosetta had the highest levels of heterozygosity of the pure seed batches (Figure 4.5 B). There were no consistent trends in Study 2 (like Study 1), with some increasing combinations more substantially from OS to S19 followed by a decline in A19. However, other seed combinations did not follow this pattern with sporadic changes between replicates of the same seed combination (Appendix 17; Appendix 18)

Table 4.5. ANOVA results of the mean H_e based on SNPs for Study 2

Variable	d.f	SS	MSS	F	P	Sig.
Replicate	2	0.000	0.000	0.137	0.872	ns
Time point	3	0.000	0.000	8.771	4.58e-05	***
Seed combination	14	0.010	0.000	28.120	< 2e-16	***
Time point:Seed combination	40	0.001	0.000	1.171	0.273	ns
Residuals	76	0.002	0.000			

***(p< 0.001); **(p< 0.01); *(p< 0.05); ns = not significant

4.4 Discussion

The practice of sowing seed mixtures in Ireland is common due to the perceived benefits of sowing cultivars with complementary characteristics. This is despite conflicting evidence demonstrating the benefits of mixed swards and the long term stability of sown PRG proportions (Pollnac *et al*., 2014; Lowry *et al*., 2020; Tubritt *et al*., 2021). It is also unclear how much each cultivar contributes to the overall performance of the sward. Determining the changes of seed mixture components over time, and their genotypic composition, is therefore critical to understanding the benefit of swards with multiple cultivars. To assess this, we employed SeqSNP to track changes in the genetic composition of complex PRG mixtures over a period

of two years with sown seed (OS) acting as a baseline for genetic diversity. SeqSNP performed well as a sequencing approach, with missing data predominantly attributed to samples rather than targets. When problematic samples were removed, much lower levels of missing data was observed (~5%). SeqSNP utilises single primer enrichment technology (SPET), a scalable and precision based amplification approach (Scaglione *et al.*, 2019). Low levels of missingness were also observed in when SPET was deployed in tomato (*Solanum lycopersicum)* and eggplant (*Solanum melongena*) (Barchi *et al*., 2019). Double probes were utilised in this study to reduce the instance of missing data and alleviate the missingness and uneven coverage encountered in the development of GenoGrass. We also increased the number of targets captured, as though few polymorphism can reliably predict traits in PRG (Byrne *et al*., 2016), moderate increases in target number potentially offsets the percentage of missing data associated target failure. Single probes can be also utilised to further scale up target number while reducing primer interactions (Barchi *et al.,* 2019). The efficiency, low cost and flexibility of SeqSNP makes it suitable for many applications in PRG genetic characterisation and breeding.

Genetic fingerprints of each of the mixtures and pure species were generated for comparison over time (Byrne *et al*., 2014; Verwimp *et al*., 2018). The genetic profiles of mixtures were generally intermediates of all mixture components. After one year post sowing the mixtures were consistently stable. The OS time points generally cluster more closely around a single component cultivar and then drift is demonstrated over time (Figure 4.1 B and Figure 4.2 B). Griffith *et al*. (2016) observed that dominant components in terms of proportionality, drifted to equal proportions. Griffith *et al*. (2016) utilised simulated grazing rather than animal grazing. This would eliminate the impact of preferential grazing which could affect the dominance of tetraploid components. In both Study 1 and Study 2 cultivars of both diploids and tetraploids (with the exception of three way mixtures) were sown in equal proportions. Though measures were undertaken to collect a representative sample, the subsampling method does not ensure that equal amounts of each seed are captured in the OS samples. Each seed component was weighed and seed combinations were then subsampled from the sown seed to observe a linear time line from sowing to sampling at each timepoint. This may have introduced

overrepresentation of single cultivars which was observed in the OS timepoints, this was particularly the case where seed size of cultivars are different (tetraploid seed is larger than diploid). Weighing equal quantities of seed for the original seed batches to establish a baseline level of genetic diversity would mitigate ambiguity of component proportionality that subsampling sown seed may introduce. Variation between mixed samples was greater than pure cultivars. Hand grabs appear to be suitable for pure cultivars, but mixtures may benefit from single leaf samples as in Verwimp *et al*. (2018). Though a two year period is not reflective of the typical life span of a sward (Shalloo *et al*., 2004), Faville *et al*. (2020) noted little change in the genetic profile of single cultivar swards over a time period of 5-6 years, using SSR markers.

In both studies, mean H_e increased from the OS timepoint to the later time points S18 and S19, followed by a decline at the time point A19 (with some exceptions). The decreases in mean H_e from the S19 to A19, could be due to herbivory of cows removing biomass over the grazing season, however this trend was not consistent. Genetic diversity of the seed combinations was observed to change over time, with the time point variable having a significant effect and OS differing significantly from all other time points in Study 2. Unsurprisingly, as the number of mixture components increased, so did the level of heterozygosity (Figure 4.4; and Figure 4.3). The Irish bred (Teagasc; AFBI) cultivars (Solas; Oakpark; Rosetta) were observed to be most heterozygous. These cultivars were also observed to dominate the mixed seed bathes though this dominance often diminished over time. This was sporadic and inconsistent across time and seed combinations.

In comparison of the agronomic performance of the seed combinations, Tubritt *et al*. (2021) investigated important agronomic features and if overyielding and synergism could be detected in complex mixtures. A single seed combination Astonenergy/Solas yielded significantly greater than the mean of its constituents. Although some studies have shown over-yielding, it has been suggested that significant diversity between components is required to produce mixed swards which outyield their mean components (Nyfeler *et al*., 2009). Tubritt *et al*. (2021) concluded that measuring characteristics in monocultures and calculating the mean yield was sufficient to accurately predict mixture performance. This supports the

trends we observed of relative stability in the genetic profile of mixtures in the post sowing time points. While evidence suggests that there is little indication of synergism and compositional changes in PRG cultivar mixtures over a 2 year period, longer term studies may be required to definitively elucidate the changes over the lifetime of a sward.

4.5 Conclusion

The evidence from this study suggests that PRG mixtures were stable over two years post sowing. There were some changes detected form original seed mixtures, however there were no consistent trends observed. Targeted low cost sequencing approaches are suitable for characterising genetic diversity in PRG mixed and pure cultivar swards.

Chapter 5
General discussion

5.1 General discussion

Forages provide a cheap source of feed for the beef and dairy industry. Perennial ryegrass (PRG) is by far the most widely sown grass species in Ireland, accounting for 95% of the grass seed sold (DAFM, 2021). The improvement of economically important traits in PRG is essential for the further intensification of beef and dairy systems in Ireland, while maintaining resilience and environmental sustainability of the sector (Parsons *et al*., 2011; O'Mara, 2012). The integration of genomic tools into crop improvement may vastly accelerate the genetic gain of such crops as well as improve our understanding of how traits are inherited and expressed (Hayes *et al*., 2013). Forages, as an annual crop, can be impacted in their composition by pressures present on farm. These pressures affect their ability to persist in the sward (Clark, 2011). As mixtures, comprised of several different cultivars, are widely sown in Ireland, differing persistency between cultivars may affect the survival of as well as their contribution to sward performance.

5.1.1 Persistency

Persistency in forage grasses is a complex trait with many contributing factors (Curran *et al*., 2020). The improvement of persistency is a breeding objective for the long term sustainable improvement of PRG (O'Donovan *et al*., 2011). Improvement of traits can be undertaken by traditional phenotypic selection or more modern techniques of marker assisted selection and genomic selection. Hybridisation of *Lolium* and *Festuca* is a novel approach to improve the persistency while maintaining the positive grazing attributes.

The trends in persistency observed in our study in Chapter 2 demonstrated that hybrid material tended to align more closely to the *Lolium* pure species controls. Values for ground score demonstrated comparability between *L. multiflorum* controls, hybrid types, and *Festuca* controls. *Lolium perenne* pure species controls exhibited greater persistency, this was expected (Wilkins & Humphreys, 2003), while the hybrid *L. perenne x F. pratensis* also displayed ground scores similar to that of its *Lolium* component. *Lolium multiflorum* controls were expected to be the least persistent, (Wilkins & Humphreys 2003), however they were found to have intermediate persistency. Due to the resilience to abiotic stress observed in *Festuca*, it was expected for these cultivars to exhibit enhanced persistency.

However, ground score cover is a measure of abundance and does not necessarily reflect productivity as demonstrated by the cultivar Fure, the *F. pratensis* pure species control. Although ground score decreases, dry matter yield increases after an extended period, perhaps due to the broad leaf morphology of *Festuca* species in comparison to *Lolium* species (Gymer and Whittington, 1972).

Despite the presence of open ground, indicated by the low ground score, the larger leaf area and subsequent larger volume of biomass may account for the disparity between ground score as a measure of persistency and the observed yield. Ground score measures the abundance of a species within a sward rather than the biomass of the surviving plants, failing to account for yield stability. Methods have been proposed such as reporting the relationship between short term and long term yield (Dodd *et al*., 2018). However, measuring yield stability captures the agronomic trait but does not explain if declines are attributed to losses in tillering ability of individual plants or losses of the individual plants themselves. This was highlighted by O'Donovan *et al*. (2017) who stated a molecular approach would be a gold standard for quantifying persistency within a sward.

5.1.2 Reduced representation assays

Reduced representation genome sequencing is suitable for applications in crop breeding due to the relatively lower cost and the potential of fewer SNPs in breeding applications (Byrne *et al*., 2017; Zhang *et al*., 2020). We intended to demonstrate the utility of such SNP arrays in population delineation in a number of population types to investigate the ability of the assay to detect proportional changes in PRG swards. Reducing genome complexity can be carried out by way of a multiplex PCR rather than an unspecific enzymatic approach (Elshire *et al*., 2011; Campbell *et al*., 2015). Genotyping in thousands by sequencing (GT-seq) is a cost effective, high throughput assay based on amplicon sequencing (Campbell *et al*., 2015). It offers a targeted approach to genotyping to interrogate a comparatively small number of regions of interest, utilising a two-fold multiplex PCR and dual barcode system. With the use of genomic resources available for PRG we have developed GenoGrass in this thesis. It utilises a similar approach to GT-seq but has been designed to target specific regions of the PRG genome for functional processes in breeding. This approach offers the potential to increase read depth at individual loci, without

increasing sequencing costs. This is beneficial in the case of high diversity species, as greater read depth aids reduction of ambiguous variant calls (Sims *et al*., 2014). A limitation encountered while developing GenoGrass was the unevenness of coverages and missingness encountered. Missingness is often overcome by imputation, however there are errors associated with missing data (O'Leary *et a*l., 2018). The incomplete nature of the PRG draft genome may hamper approaches such as these (Byrne *et al*., 2015; Veeckman *et al*., 2016) The improvement of genomic resources in PRG is ongoing (Roldan Ruiz *et al*., 2019). For further applications we attempted to use an alternative approach, SeqSNP, which offered greater specificity while retaining the benefits of this type of assay.

SeqSNP was then implemented to investigate changes in PRG mixtures. This performed well as a sequencing approach, with missing data predominantly attributed to samples rather than targets. When problematic samples were removed, much lower levels of missing data was observed (~5%). SeqSNP was found to provide adequate depth of coverage and even representation across the selected target regions, allowing for the characterisation of genetic diversity over time. SeqSNP utilises single primer enrichment technology (SPET), a scalable and precision based amplification approach (Scaglione *et al*., 2019). GT-seq and SeqSNP captures novel variants as well as target SNPs (Campbell et al., 2015). As it captures both target and novel variants there is future potential for integrating multiple SNPs this into haplotypes to improve discriminatory power. SNPs are the smallest unit of genetic variation and when treated independently can reveal a limited amount of information. Combining a small number of SNPs into haplotypes may increase the informative power relative to a single, independent SNP. The integration of SNPs into haplotypes was found to have the potential to improve accuracy in population assignment (Gattepaille & Jakobsson, 2012). The combination of fewer more powerful mutilallelic markers can be integrated as an approach into future projects.

Reduced representation assays of this kind have wider applications in forage breeding and cultivar evaluation. Increasing workloads and costs have been identified as the principal prospective issues in the Distinctness, Uniformity and Stability (DUS) system (Gilliland & Gensollen, 2010). The incorporation of molecular

markers into the DUS tests can enhance the protection process and improve certification efficiency (Wang *et al*., 2016). DUS acts as a protection system for plant breeders rights. To attain protected status, it must be proven that the variety is distinct from all existing varieties in “common knowledge”. Guidelines for evaluation are set out by the International Union for the Protection of New Varieties of Plants (UPOV), and once “distinctiveness” is confirmed, the variety can be certified. Certification is awarded through the EU agency, the Community Plant Variety Office (CPVO, Angers, France). Certification is valid in all EU member states. For a plant variety to be added to National Lists, it must be certified under this process. This evaluation is carried out by each member under a national body. Aside from ploidy, all characteristics evaluated are morphological. A grow out test is required to asses vegetative and reproductive characteristics of PRG, and comparisons are made to a large reference collection. The characters observed in this evaluation can be effected by environmental factors and in allogamous populations such as PRG, it can be difficult to conduct tests while mitigating variability. The benefit of incorporating molecular techniques in DUS tests has been recognised by UPOV, in particular to aid the management of the reference collection and reducing variability in testing (UPOV, 2014). Reliable, reproducible markers which are cost effective are required for such systems. Low cost assays which utilise SNPs have been identified as promising options for the incorporation into protocols (Gilliland *et al*., 2019).

5.1.2 Mixtures

The dynamic changes in sward structure over time can be observed in cultivars of mixed swards. Differences in ability to persist of individual plants as well as cultivars has affected the mixtures. Although some studies have shown over-yielding, it has been suggested that significant diversity between components is required to produce mixed swards which out yield their mean components (Nyfeler *et al*., 2009). Morphological similarities between cultivars results in difficulties when discerning changes in swards sown with several cultivars. For this reason it is useful to employ a molecular technique for genotype differentiation. Previous Teagasc research investigated the change of sward composition over time by examining phosphoglucoisomerase allozyme frequencies relative to the frequencies of the component cultivars in binary mixtures (one tetraploid and one diploid) under

simulated grazing (Griffith *et al*., 2016). This study found that despite varying sowing ratios, established sward composition would differ from sown composition when competitive advantage was possessed by a single cultivar. We hypothesised that complex mixtures, reflective of mixtures sown with contrasting characteristics would change over time. When we investigated changes in genetic composition, some shifts in genetic profiles were observed from the original seed, however these changes were sporadic and no clear trends or relationships were observed. Little change was also observed in pure and binary PRG swards (Verwimp *et al*. 2018; Faville *et al*., 2020). Both time point and seed combination were found to have a significant effect on genetic diversity (mean H_e) however, only a in Study 2 was a time point found to differ significantly from another.

5.2 Conclusion

Reduced representation assays utilising SNPs have the potential to accelerate the improvement and understanding of persistency and sward dynamics in PRG. We have demonstrated the limitations of current methods of quantification of persistency. We attempted to utilise low cost and efficient method of characterising genetic diversity in a range of PRG populations. Comparatively fewer SNPs are required to adequately delineate between PRG populations. A revised SNP panel was used to track changes in complex PRG mixtures. Little compositional changes in were observed over a two year period. Future work should continue to monitor the changes in complex PRG mixtures over an extended period.

References

Aird, D., Ross, M.G., Chen, W.S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum, C. and Gnirke, A., 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. Genome Biology, 12(2), pp.1-14.

Anhalt, U.C.M., Pat, J.H.H., Piepho, H.P., Byrne, S. and Barth, S., 2009. Quantitative trait loci mapping for biomass yield traits in a *Lolium* inbred line derived F 2 population. Euphytica, 170(1-2), pp.99-107.

Arojju, S.K., Barth, S., Milbourne, D., Conaghan, P., Velmurugan, J., Hodkinson, T.R. and Byrne, S.L., 2016. Markers associated with heading and aftermath heading in perennial ryegrass full-sib families. BMC Plant Biology, 16(1), p.160.

Arojju, S.K., Conaghan, P., Barth, S., Milbourne, D., Casler, M.D., Hodkinson, T.R., Michel, T. and Byrne, S.L., 2018. Genomic prediction of crown rust resistance in *Lolium perenne*. BMC Genetics, 19(1), pp.1-10.

Arojju, S.K., Cao, M., Trolove, M., Barrett, B.A., Inch, C., Eady, C., Stewart, A. and Faville, M.J., 2020. Multi-trait genomic prediction improves predictive ability for dry matter yield and water-soluble carbohydrates in perennial ryegrass. Frontiers in Plant Science, 11, 1197.

Banzhaf, K. and Boberfeld, W.O.V., 2005, June. Ensilability and silage quality of different Festulolium hybrids in comparison to *Festuca arundinacea*. In XX International Grassland Conference: Offered papers: Offered Papers (p. 472). Wageningen Academic Publishers.

Barth, S., McGrath, S.K., Arojju, S.K. and Hodkinson, T.R., 2017. An Irish perennial ryegrass genetic resource collection clearly divides into two major gene pools. Plant Genetic Resources, 15(3), p.269.

Bevan, M.W., Uauy, C., Wulff, B.B., Zhou, J., Krasileva, K. and Clark, M.D., 2017. Genomic innovation for crop improvement. Nature, 543(7645), pp.346-354.

Blackmore, T., Thomas, I., McMahon, R., Powell, W. and Hegarty, M., 2015. Genetic–geographic correlation revealed across a broad European ecotypic sample of perennial ryegrass (*Lolium perenne*) using array-based SNP genotyping. Theoretical and Applied Genetics, 128(10), pp.1917-1932.

Bolaric, S., Barth, S., Melchinger, A.E. and Posselt, U.K., 2005. Molecular characterization of genetic diversity in European germplasm of perennial ryegrass. Euphytica, 146(1), pp.39-44.

Byrne, S., Czaban, A., Studer, B., Panitz, F., Bendixen, C. and Asp, T., 2013. Genome wide allele frequency fingerprints (GWAFFs) of populations via genotyping by sequencing. Plos One, 8(3), p.e57438.

Byrne, S.L., Conaghan, P., Barth, S., Arojju, S.K., Casler, M., Michel, T., Velmurugan, J. and Milbourne, D., 2017. Using variable importance measures to identify a small set of SNPs to predict heading date in perennial ryegrass. Scientific Reports, 7(1), pp.1-10.

Byrne, S.L., Nagy, I., Pfeifer, M., Armstead, I., Swain, S., Studer, B., Mayer, K., Campbell, J.D., Czaban, A., Hentrup, S. and Panitz, F., 2015. A synteny-based draft genome

sequence of the forage grass *Lolium perenne*. The Plant Journal, 84(4), pp.816-826.

Campbell, N.R., Harmon, S.A. and Narum, S.R., 2015. Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. Molecular Ecology Resources, 15(4), pp.855-867.

Cashman, P.A., Gilliland, T.J., McEvoy, M., Watson, S. and O'Donovan, M., 2016. Changes in plant morphological expression in 12 perennial ryegrass cultivars following frequent and infrequent cutting management. The Journal of Agricultural Science, 154(03), pp.456-471.

Casler, M.D., Peterson, P.R., Hoffman, L.D., Ehlke, N.J., Brummer, E.C., Hansen, J.L., Mlynarek, M.J., Sulc, M.R., Henning, J.C., Undersander, D.J. and Pitts, P.G., 2002. Natural selection for survival improves freezing tolerance, forage yield, and persistence of *Festulolium*. Crop Science, 42(5), pp.1421-1426.

Castle, M.E. and Watson, J.N., 1971. A comparison between a diploid and a tetraploid ryegrass for milk production. The Journal of Agricultural Science, 77(1), pp.69-76.

Clark, D.A., 2011. Changes in pastoral farming practices and pasture persistence–a review. In Pasture persistence symposium.

Conaghan, P. and Casler, M.D., 2011. A theoretical and practical analysis of the optimum breeding system for perennial ryegrass. Irish Journal of Agricultural and Food Research, 50, pp.47-63.

Cowling, D.W. and Lockyer, D.R., 1968. A comparison of the yield of three grass species at various levels of nitrogenous fertilizer sown alone or in a mixture. The Journal of Agricultural Science, 71(1), pp.127-136.

Creighton, P., Kennedy, E., Shalloo, L., Boland, T.M. and O'Donovan, M., 2011. A survey analysis of grassland dairy farming in Ireland, investigating grassland management, technology adoption and sward renewal. Grass and Forage Science, 66(2), pp.251-264.

Culleton, N., Cullen, T. and McCarthy, V., 1992. The decline of the herbage seed production industry in Ireland. Irish Geography, 25(1), pp.98-101.

Curran, N., Grogan, D., Milbourne, D., Byrne, S.L., O'Riordan, E., Hanley, M., Grant, T, Hodkinson, T.R., Barth, S., 2020. Persistency, yield, and silage quality of Festulolium cultivars over a consecutive five-year period under a mild Atlantic climate. Biologia Plantarum, 64, pp.856-864.

DAFM, 2021. Grass and White Clover Recommended List Varieties for Ireland 2021. Publication produced by From Department of Agriculture, Food and the Marine, Ireland. Available: https://www.gov.ie/en/publication/8dbf3-grass-and-white-clover-recommended-list-varieties-for-ireland-2021/

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T. and McVean, G., 2011. The variant call format and VCFtools. Bioinformatics, 27(15), pp.2156-2158.

Davies, D.R., Merry, R.J., Williams, A.P., Bakewell, E.L., Leemans, D.K. and Tweed, J.K.S., 1998. Proteolysis during ensilage of forages varying in soluble sugar content. Journal of Dairy Science, 81(2), pp.444-453.

Diekmann, K., Hodkinson, T.R. and Barth, S., 2012. New chloroplast microsatellite markers suitable for assessing genetic diversity of *Lolium perenne* and other related grass species. Annals of Botany, 110(6), pp.1327-1339.

Dillon, P., Hanrahan, L., Thorne, F., Kelly, P., Shalloo, L. and Moran, B., Irish dairying: well positioned for the upturn. In Teagasc National Dairy Conference (p. 8).

Dodd, M.B., Chapman, D.F., Ludemann, C.I., Griffiths, W., Tozer, K.N. and Donnelly, L., 2018. The measurement of perennial ryegrass persistence. Journal of New Zealand Grasslands, 80, pp.161-168.

Doyle, J., 1991. DNA protocols for plants. In Molecular techniques in taxonomy (pp. 283-293). Springer, Berlin, Heidelberg.

Duchini PG, Guzatti GC, Echeverria JR, Américo LF, Sbrissia AF. Experimental evidence that the perennial grass persistence pathway is linked to plant growth strategy. PloS one. 2018 Nov 26;13(11):e0207360.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. and Mitchell, S.E., 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS one, 6(5), p.e19379.

Faville, M.J., Crush, J.R., Hong, W., Phillips, H., Lee, J.M. and Chapman, D., 2020. Effects of pasture age on the genotype and phenotype of perennial ryegrass. Grass and Forage Science, 75(2), pp.135-144.

Finneran, E., Crosson, P., O'kiely, P., Shalloo, L., Forristal, D. and Wallace, M., 2010. Simulation modelling of the cost of producing and utilising feeds for ruminants on Irish farms. Journal of Farm Management, 14(2), pp.95-116.

Foito, A., Hackett, C.A., Byrne, S.L., Stewart, D. and Barth, S., 2015. Quantitative trait loci analysis to study the genetic regulation of non-polar metabolites in perennial ryegrass. Metabolomics, 11(2), pp.412-424.

Fojtik, A.: Methods of grass improvement used at the Plant Breeding Station Hladke Zivotice. – Genetica Polonica, 35: 25-31, 1994.

Fox, J. and Weisberg, S., 2018. An R companion to applied regression. Thousand Oaks, CA: Sage publications.

Ghesquière, M., Baert, J., Barth, S., Černoch, V., Grogan, D., Humphreys, M.W., Murray, P., Østrem, L., Sokolović, D., Paszkowski, E. and Zwierzykowski, Z., 2016. Enhancing the productivity in forage grasses on the European scale using interspecific hybridization. Breeding in a world of scarcity. In: Proceedings of the 2015 meeting of the section ‘‘Forage Crops and Amenity Grasses’’ of Eucarpia. Springer International Publishing, Ghent ,Belgium pp 199–204

Ghesquière, M., Emile, J.C., Jadas-Hécart, J., Mousset, C., Traineau, R. and Poisson, C., 1996. First in vivo assessment of feeding value of Festulolium hybrids derived from Festuca arundinacea var. glaucescens and selection for palatability. Plant Breeding, 115(4), pp.238-244.

Gilliland, T.J., Johnston, J. and Connolly, C., 2007. A review of forage grass and clover seed use in Northern Ireland, UK between 1980 and 2004. Grass and Forage Science, 62(3), pp.239-254.

Gilliland, T.J. and Gensollen, V., 2010. Review of the protocols used for assessment of DUS and VCU in Europe–Perspectives. In Sustainable Use of Genetic Diversity in Forage and Turf Breeding (pp. 261-275). Springer Netherlands.

Gilliland, T.J., Hennessy, D. and Griffith, V., 2011. Studies into the dynamics of perennial ryegrass (*Lolium perenne* L.) seed mixtures. Irish Journal of Agricultural and Food Research, pp.99-112.

Gilliland, T.J., Annicchiarico, P., Julier, B. and Ghesquière, M., 2019. Revising official herbage cultivar evaluation to meet evolving EU stakeholder needs. Improving sown grasslands through breeding and management 24 (2019): p.475.

Griffith, V., Gilliland, T.J., McEvoy, M. and O'Donovan, M., 2016. Competitive dynamics in mixtures of perennial ryegrass cultivars during two growing seasons under two cutting managements. Crop and Pasture Science, 67(2), pp.225-235.

Grogan, D. and Gilliland, T.J., 2011. A review of perennial ryegrass variety evaluation in Ireland. 50 (1), Irish Journal of Agricultural and Food Research, pp.65-81.

Grogan, D., Barth, S., Grant, J., O'Riordan, E. and Hanley, M., 2018. *Festulolium* in Ireland–seasonal yield and quality assessment. Sustainable meat and milk production from grasslands, p.206.

Halling, M.A., 2012. Yield stability of *Festulolium* and perennial ryegrass in southern and central Sweden. *Grassland Science in Europe*, *17*, pp.118-120.

Hayes, B.J., Cogan, N.O., Pembleton, L.W., Goddard, M.E., Wang, J., Spangenberg, G.C. and Forster, J.W., 2013. Prospects for genomic selection in forage plant species. Plant Breeding, 132(2), pp.133-143.

Houdek, I., 2005. Festulolium 'Perseus'. Czech Journal of Genetics and Plant Breeding, 41, pp.35-36.

Huang, B. and Gao, H., 2000. Root physiological characteristics associated with drought resistance in tall fescue cultivars. Crop Science, 40(1), pp.196-203.

Humphreys, J. and Casey, I.A., 2002, April. Grassland renovation in Ireland. In Grassland Resowing and Grass-Arable Rotations. Proceedings of European Grassland Federation International Workshop Agricultural and Environmental Issues, Wageningen (The Netherlands), (Vol. 18, pp. 79-91).

Humphreys, M.W., Canter, P.J. and Thomas, H.M., 2003. Advances in introgression technologies for precision breeding within the *Lolium-Festuca* complex. Annals of Applied Biology, *143*(1), pp.1-10.

Humphreys, M.W., O'Donovan, S.A., Farrell, M.S., Gay, A.P. and Kingston-Smith, A.H., 2014. The potential of novel *Festulolium* (2n= 4x= 28) hybrids as productive, nutrient-use-efficient fodder for ruminants. Food and Energy Security, *3*(2), pp.98-110.

Humphreys, M.W. and Zwierzykowski, Z., 2020. *Festulolium*, a century of research and breeding and its increased relevance in meeting the requirements for multifunctional grassland agriculture. Festulolium: From the Nature to Modern Breeding, p.1.

Hurtado-Uria, C., Hennessy, D., Shalloo, L., O'Connor, D. and Delaby, L., 2013. Relationships between meteorological data and grass growth over time in the south of Ireland. Irish Geography, 46(3), pp.175-201.

Jombart, T., 2008. adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics, 24(11), pp.1403-1405.

Jombart, T., Devillard, S. and Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics, 11(1), p.1-15.

Jones, E.L. and Roberts, J.E., 1994. Herbage quality and production of perennial ryegrass cultivars in monoculture and mixtures. Irish Journal of Agricultural and Food Research, 33(2) pp.169-176.

Jones, N., Ougham, H., Thomas, H. and Pašakinskienė, I., 2009. Markers and mapping revisited: finding your gene. New Phytologist, 183(4), pp.935-966.

Kelly, P., Shalloo, L., Wallace, M. and Dillon, P., 2020. The Irish dairy industry–Recent history and strategy, current state and future challenges. International Journal of Dairy Technology, 73(2), pp.309-323.

Kopecký, D., Šimoníková, D., Ghesquière, M. and Doležel, J., 2017. Stability of genome composition and recombination between homoeologous chromosomes in *Festulolium* (Festuca× Lolium) cultivars. Cytogenetic and Genome Research, *151*(2), pp.106-114.

Lemežienė, N., Kanapeckas, J., Tarakanovas, P. and Nekrošas, S., 2004. Analysis of dry matter yield structure of forage grasses. Plant Soil Environment, *50*(6), pp.277-282.

Lenth, R., Singmann, H., Love, J., Buerkner, P. and Herve, M., 2019. Estimated marginal means, aka least-squares means. R package version 1.3. 2.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R., 2009. The sequence alignment/map format and SAMtools. Bioinformatics, 25(16), pp.2078-2079.

Li, H. and Durbin, R., 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics, 26(5), pp.589-595.

Lin, B.B., 2011. Resilience in agriculture through crop diversification: adaptive management for environmental change. BioScience, 61(3), pp.183-193.

Linck, E. and Battey, C.J., 2019. Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. Molecular Ecology Resources, 19(3), pp.639-647.

Lowry, C.J., Bosworth, S.C., Goslee, S.C., Kersbergen, R.J., Pollnac, F.W., Skinner, R.H., Warren, N.D. and Smith, R.G., 2020. Effects of expanding functional trait diversity on productivity and stability in cultivar mixtures of perennial ryegrass. Agriculture, Ecosystems & Environment, 287, p.106691.

Lynch, T.M., Barth, S., Dix, P.J., Grogan, D., Grant, J. and Grant, O.M., 2015. Ground cover assessment of perennial ryegrass using digital imaging. Agronomy Journal, 107(6), pp.2347-2352.

Macleod, C., Humphreys, M.W., Whalley, W.R., Turner, L., Binley, A., Watts, C.W., Skøt, L., Joynes, A., Hawkins, S., King, I.P., O'Donovan, S. and Haygarth, P.M., 2013. A novel grass hybrid to reduce flood generation in temperate regions. Scientific reports, 3(1), pp.1-7.

Mayne, C.S. and O'Kiely, P., 2005, July. An overview of silage production and utilisation in Ireland (1950–2005). In Proceedings of the XIV international silage conference. Wageningen Academic Publishers, Belfast, North Ireland (pp. 19-34).

McDonagh, J., O'Donovan, M., McEvoy, M. and Gilliland, T.J., 2016. Genetic gain in perennial ryegrass (*Lolium perenn*e) varieties 1973 to 2013. Euphytica, 212(2), pp.187-199.

McEvoy, M., O'Donovan, M. and Shalloo, L., 2011. Development and application of an economic ranking index for perennial ryegrass cultivars. Journal of Dairy Science, 94(3), pp.1627-1639.

McGrath, S., Hodkinson, T.R. and Barth, S., 2007. Extremely high cytoplasmic diversity in natural and breeding populations of *Lolium* (Poaceae). Heredity, 99(5), pp.531-544.

McGrath, S., Hodkinson, T.R., Charles, T.M., Zen, D.G. and Barth, S., 2010. Variation in inflorescence characters and inflorescence development in ecotypes and cultivars of Lolium perenne L. Grass and Forage Science, 65(4), pp.398-409

McGrath, S., Hodkinson, T.R., Frohlich, A., Grant, J. and Barth, S., 2014. Seasonal and genetic variations in water-soluble carbohydrates and other quality traits in ecotypes and cultivars of perennial ryegrass (*Lolium perenne* L.). Plant Genetic Resources, 12(2), pp.236-247.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. and DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research, 20(9), pp.1297-1303.

Meehan, P., Burke, B., Doyle, D., Barth, S. and Finnan, J., 2017. Exploring the potential of grass feedstock from marginal land in Ireland: Does marginal mean lower yield?. Biomass and Bioenergy, 107, pp.361-369.

Muck, R.E., 1988. Factors influencing silage quality and their implications for management. Journal of Dairy Science, 71(11), pp.2992-3002.

Muhandiram, N.P., Humphreys, M.W., Fychan, R., Davies, J.W., Sanderson, R. and Marley, C.L., 2020. Do agricultural grasses bred for improved root systems provide resilience to machinery-derived soil compaction?. Food and Energy Security, 9(3), p.e227.

Myers, W.M., 1939. Colchicine induced tetraploidy in perennial ryegrass: *Lolium perenne* L. Journal of Heredity, 30(11), pp.499-504.

Neuteboom, J.H., Lantinga, E.A. and Wind, K., 1988, July. Tillering characteristics of diploid and tetraploid perennial ryegrass. Proceedings of the 12th General Meeting of the European Grassland Federation, Dublin. Irish Grassland Association, Belclare, pp. 498–503.

Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E., Connolly, J. and Lüscher, A., 2009. Strong mixture effects among four species in fertilized agricultural grassland led to persistent and consistent transgressive overyielding. Journal of Applied Ecology, 46(3), pp.683-691.

O'Donovan M., McHugh, N., McEvoy, M., Grogan, D. and Shalloo, L., 2017. Combining seasonal yield, silage dry matter yield, quality and persistency in an economic index to assist perennial ryegrass variety selection. The Journal of Agricultural Science, 155(4), pp.556-568.

O'Donovan, M., Lewis, E. and O'Kiely, P., 2011. Requirements of future grass-based ruminant production systems in Ireland. Irish Journal of Agricultural and Food Research, 50 (1), pp.1-21.

O'Leary, S.J., Puritz, J.B., Willis, S.C., Hollenbeck, C.M. and Portnoy, D.S., 2018. These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. Molecular Ecology, 27 (16), pp. 3193-3206

O'Mara, F.P., 2012. The role of grasslands in food security and climate change. Annals of Botany, 110(6), pp.1263-1270.

Ogiso-Tanaka, E., Shimizu, T., Hajika, M., Kaga, A. and Ishimoto, M., 2019. Highly multiplexed AmpliSeq technology identifies novel variation of flowering time-related genes in soybean (*Glycine max*). DNA Research, 26(3), pp.243-260.

Østrem, L., Volden, B. and Larsen, A., 2013. Morphology, dry matter yield and phenological characters at different maturity stages of× *Festulolium* compared with other grass species. Acta Agriculturae Scandinavica, Section B-Soil & Plant Science, 63(6), pp.531-542.

Parsons, A.J., Edwards, G.R., Newton, P.C.D., Chapman, D.F., Caradus, J.R., Rasmussen, S. and Rowarth, J.S., 2011. Past lessons and future prospects: plant breeding for yield and persistence in cool-temperate pastures. Grass and Forage Science, 66(2), pp.153-172.

Pembleton, L.W., Drayton, M.C., Bain, M., Baillie, R.C., Inch, C., Spangenberg, G.C., Wang, J., Forster, J.W. and Cogan, N.O., 2016. Targeted genotyping-by-sequencing permits cost-effective identification and discrimination of pasture grass species and cultivars. Theoretical and Applied Genetics, 129(5), pp.991-1005.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and Team, R.C., 2007. Linear and nonlinear mixed effects models. R package version, 3(57), pp.1-89.

Playne, M.J. and McDonald, P., 1966. The buffering constituents of herbage and of silage. Journal of the Science of Food and Agriculture, 17(6), pp.264-268.

Pollnac, F.W., Smith, R.G. and Warren, N.D., 2014. Cultivar diversity as a means of ecologically intensifying dry matter production in a perennial forage stand. Ecosphere, 5(9), pp.1-12.

Quaite, E. and Camlin, M.S., 1986. The use of electrophoretically labelled populations for intra-specific competition studies in perennial ryegrass (Lolium perenne L.) swards. The Journal of Agricultural Science, 106(1), pp.99-106.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Riday, H., 2011. Paternity testing: A non-linkage based marker-assisted selection scheme for outbred forage species. Crop Science, 51(2), pp.631-641.

Roldán-Ruiz, I., Ruttink, T., Muylle, H., Verwimp, C., Veeckman, E., Cnops, G., Vleugels, T. and Barre, P., 2019. Advances and perspectives in the application of genomic tools for research and breeding of forage crops. Research Collection, p.229.

Rognli, O.A., Saha, M.C., Bhamidimarri, S. and van der Heijden, S., 2010. Fescues. In Fodder crops and amenity grasses (pp. 261-292). Springer, New York, NY.

Roudier, P., Andersson, J.C., Donnelly, C., Feyen, L., Greuell, W. and Ludwig, F., 2016. Projections of future floods and hydrological droughts in Europe under a+ 2 C global warming. Climatic change, 135(2), pp.341-355.

Scaglione, D., Pinosio, S., Marroni, F., Di Centa, E., Fornasiero, A., Magris, G., Scalabrin, S., Cattonaro, F., Taylor, G. and Morgante, M., 2019. Single primer enrichment technology as a tool for massive genotyping: a benchmark on black poplar and maize. Annals of Botany, 124(4), pp.543-551

Shalloo, L., Dillon, P., O'loughlin, J., Rath, M. and Wallace, M., 2004. Comparison of a pasture-based system of milk production on a high rainfall, heavy-clay soil with that on a lower rainfall, free-draining soil. Grass and Forage Science, 59(2), pp.157-168.

Shalloo, L., Creighton, P. and O'Donovan, M., 2011. The economics of reseeding on a dairy farm. Irish Journal of Agricultural and Food Research, 50(1) pp.113-122.

Sims, D., Sudbery, I., Ilott, N.E., Heger, A. and Ponting, C.P., 2014. Sequencing depth and coverage: key considerations in genomic analyses. Nature Reviews Genetics, 15(2), pp.121-132.

Solati, Z., Manevski, K., Jørgensen, U., Labouriau, R., Shahbazi, S. and Lærke, P.E., 2018. Crude protein yield and theoretical extractable true protein of potential biorefinery feedstocks. Industrial Crops and Products, 115, pp.214-226.

Stapleton, J. and Jones, M.B., 1987. Effects of vernalization on the subsequent rates of leaf extension and photosynthesis of perennial ryegrass (*Lolium perenne* L.). Grass and Forage Science, 42(1), pp.27-31.

Tallowin, J.R.B., Brookman, S.K.E. and Santos, G.L., 1995. Leaf growth and utilization in four grass species under steady state continuous grazing. The Journal of Agricultural Science, 124(3), pp.403-417.

Thomas, H. and Humphreys, M.O., 1991. Progress and potential of interspecific hybrids of Lolium and Festuca. The Journal of Agricultural Science, 117(1), pp.1-8.

Thomas, H.M., Morgan, W.G. and Humphreys, M.W., 2003. Designing grasses with a future–combining the attributes of *Lolium* and *Festuca*. Euphytica, 133(1), pp.19-26.

Tilley, J.M.A. and Terry, D.R., 1963. A two-stage technique for the in vitro digestion of forage crops. Grass and fForage science, 18(2), pp.104-111.

Tomaszewski, C., Byrne, S.L., Foito, A., Kildea, S., Kopecký, D., Doležel, J., Heslop-Harrison, J.S., Stewart, D. and Barth, S., 2012. Genetic linkage mapping in an F2 perennial ryegrass population using DArT markers. Plant Breeding, 131(2), pp.345-349.

Tozer, K.N., Chapman, D.F., Bell, N.L., Crush, J.R., King, W.M., Rennie, G.M., Wilson, D.J., Mapp, N.R., Rossi, L., Aalders, L.T. and Cameron, C.A., 2014. Botanical survey of perennial ryegrass-based dairy pastures in three regions of New Zealand: implications for ryegrass persistence. New Zealand Journal of Agricultural Research, 57(1), pp.14-29.

Tubritt, T., Delaby, L., Gilliland, T. and O'Donovan, M., 2020. An investigation into the grazing efficiency of perennial ryegrass varieties. Grass and Forage Science, 75(3), pp.253-265.

Tubritt, T., Delaby, L., Gilliland, T.J. and O'Donovan, M., 2021. The agronomic performance of perennial ryegrass varieties evaluated as monocultures and mixtures of varying complexity under animal grazing. European Journal of Agronomy, 124, p.126226.

Van Dijk, E.L., Jaszczyszyn, Y. and Thermes, C., 2014. Library preparation methods for next-generation sequencing: tone down the bias. Experimental Cell Research, 322(1), pp.12-20.

Veeckman, E., Ruttink, T. and Vandepoele, K., 2016. Are we there yet? Reliably estimating the completeness of plant genome sequences. The Plant Cell, 28(8), pp.1759-1768.

Veeckman, E., Van Glabeke, S., Haegeman, A., Muylle, H., Van Parijs, F.R., Byrne, S.L., Asp, T., Studer, B., Rohde, A., Roldán-Ruiz, I. and Vandepoele, K., 2019. Overcoming challenges in variant calling: exploring sequence diversity in candidate genes for plant development in perennial ryegrass (*Lolium perenne*). DNA Research, 26(1), pp.1-12.

Velmurugan, J., Mollison, E., Barth, S., Marshall, D., Milne, L., Creevey, C.J., Lynch, B., Meally, H., McCabe, M. and Milbourne, D., 2016. An ultra-high density genetic linkage map of perennial ryegrass (*Lolium perenne*) using genotyping by sequencing (GBS) based on a reference shotgun genome assembly. Annals of Botany, 118(1), pp.71-87.

Verwimp, C., Ruttink, T., Muylle, H., Van Glabeke, S., Cnops, G., Quataert, P., Honnay, O. and Roldan-Ruiz, I., 2018. Temporal changes in genetic diversity and forage yield of perennial ryegrass in monoculture and in combination with red clover in swards. PloS one, 13(11), p.e0206571.

Walters, R.J.K. and Evans, E.M., 1973. Annual Report of the Welsh Plant Breeding Station for 1972. Annual report - Welsh Plant Breeding Station, University College of Wales, Aberystwyth

Wilkins, P.W., 1991. Breeding perennial ryegrass for agriculture. Euphytica, 52(3), pp.201-214.

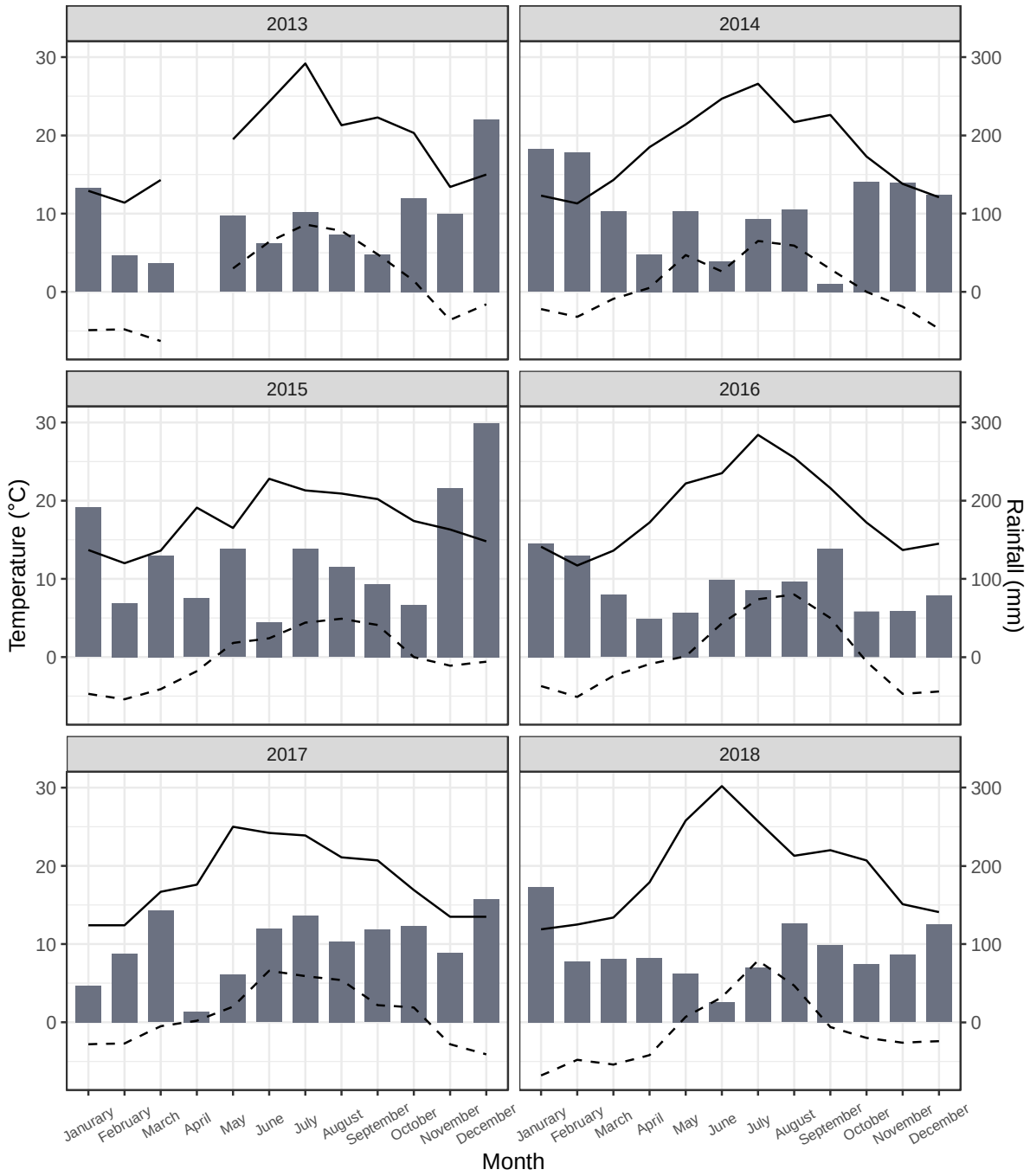
Wilkins, P.W. and Humphreys, M.O., 2003. Progress in breeding perennial forage grasses for temperate agriculture. The Journal of Agricultural Science, 140(2), pp.129-150.

Xu, Y., Li, P., Zou, C., Lu, Y., Xie, C., Zhang, X., Prasanna, B.M. and Olsen, M.S., 2017. Enhancing genetic gain in the era of molecular breeding. Journal of Experimental Botany, 68(11), pp.2641-2666.

Zhang, J., Yang, J., Zhang, L., Luo, J., Zhao, H., Zhang, J. and Wen, C., 2020. A new SNP genotyping technology Target SNP-seq and its application in genetic analysis of cucumber varieties. Scientific Reports, 10(1), pp.1-11.

Zwierzykowski, Z., Tayyar, R., Brunell, M. and Lukaszewski, A.J., 1998. Genome recombination in intergeneric hybrids between tetraploid *Festuca pratensis* and *Lolium multiflorum*. Journal of Heredity, 89(4), pp.324-328.

Appendices



Appendix 1 Meteorological data for experimental years, including total rainfall (grey bars) maximum temperature (solid line) and minimum (dashed line) for each month. April 2013 values missing for this site.

Appendix 2. Results of ANOVA, investigating the effect of year and cultivar for traits dry matter yield (DMY) and ground cover score (GC) (over five years), and quality parameters including water soluble carbohydrates (WSC), buffering capacity (BC), protein content (CP) and dry matter digestibility DMD (over 3 years).

	N	Sum Sq	DF	F value	Pr(>F)	Sig.
GC						
Block	3	3.2	2	4.205	0.0162	*
Year	5	111.9	4	88.4499	< 2.2e^{-16}	***
Cultivar	21	233	20	42.5648	< 2.2e^{-16}	***
Year : Cultivar	26	54.4	80	2.6468	1.43E-08	***
DMY						
Block	3	10	2	2.2049	0.1128	ns
Year	5	453	4	94.0058	< 2.2e-16	***
Cultivar	21	202	20	7.7019	2.86e-16	***
Year : Cultivar	26	146	80	1.3137	0.0643	ns
DMD						
Block	3	4232	2	4.54e^{+00}	0.0125	*
Year	3	146369	2	2.43e^{+02}	< 2.2e^{-16}	***
Cultivar	21	49049	20	6.28e^{+00}	2.58e^{-11}	***
Year : Cultivar	24	9960	40	1.01e^{+00}	0.4658	ns
WSC						
Block	3	3465	2	1.9355	0.148688	ns
Year	3	53554	2	58.5692	< 2.2e^{-16}	***
Cultivar	21	240788	20	16.5316	< 2.2e-16	***
Year : Cultivar	24	29674	40	2.0323	0.001627	**
BC						
Block	3	1605	2	0.3363	0.715	ns
Year	3	156132	2	62.2085	< 2.2e^{-16}	***
Cultivar	21	112698	20	4.289	2.07e^{-07}	***
Year : Cultivar	24	40290	40	0.8692	0.6888	ns
CP						
Block	3	2452	2	8.7232	0.0002854	***
Year	3	36854	2	218.6549	< 2.2e^{-16}	***
Cultivar	21	24879	20	11.9325	< 2.2e^{-16}	***
Year : Cultivar	24	4491	40	1.468	0.0572	ns

***(p< 0.001); **(p< 0.01); *(p< 0.05); ns = not significant

Appendix 3. Results of ANOVA, investigating the effect of year and hybrid type for traits dry matter yield (DMY) and ground score (GC) (over five years), and quality parameters including water soluble carbohydrates (WSC), buffering capacity (BC), protein content (CP) and dry matter digestibility DMD (over 3 years).

	N	Sum Sq	DF	F value	Pr(>F)	Sig
GC						
Block	3	3.2	2	2.126	0.1213	

Year	5	111.9	4	36.9569	$< 2.2e^{-16}$	***
Control/Hybrid type	10	233	9	34.1947	$< 2.2e^{-16}$	***
Year:Control/Hybrid type	15	54.4	36	1.9963	0.0011	**
DMY						
Block	3	10	2	1.863	0.1572	ns
Year	5	453	4	42.9865	$< 2.2e^{-16}$	***
Control/Hybrid type	10	202	9	8.4908	$3.83e^{-11}$	***
Year:Control/Hybrid type	15	146	36	1.5355	0.0314	*
DMD						
Block	3	4232	2	$4.35e^{+00}$	0.0144	*
Year	3	146369	2	$1.51e^{+02}$	$< 2.2e^{-16}$	***
Control/Hybrid type	10	49049	9	$1.12e^{+01}$	$1.95e^{-13}$	***
Year:Control/Hybrid type	13	9960	18	$1.14e^{+00}$	0.3204	ns
WSC						
Block	3	3465	2	1.3001	0.2754	
Year	3	53554	2	20.0953	$1.70e^{-08}$	***
Control/Hybrid type	10	240788	9	20.0782	$< 2.2e^{-16}$	***
Year:Control/Hybrid type	13	29674	18	1.2372	0.2381	ns
BC						
Block	3	1605	2	0.2927	0.7467	ns
Year	3	156132	2	28.4754	$2.81e^{-11}$	***
Control/Hybrid type	10	112698	9	4.5675	$2.32e^{-05}$	***
Year:Control/Hybrid type	13	40290	18	0.8165	0.6789	ns
CP						
Block	3	2452	2	6.4483	0.0020	**
Year	3	36854	2	96.9226	$< 2.2e^{-16}$	***
Control/Hybrid type	10	24879	9	14.5398	$< 2.2e^{-16}$	***
Year:Control/Hybrid type	13	4491	18	1.3123	0.1867	ns

***(p< 0.001); **(p< 0.01); *(p< 0.05); ns = not significant

Appendix 4. The mean ground cover (GC) visual score on a scale from 0 to 9 (0 = no ground cover, 9 = full ground cover) per cultivar for each year and the overall mean for a five year period, including mean yield per hybrid type for each year and five year mean. Superscripts indicate significant differences in mean GC score among cultivars.

Cultivar	**2014 GC**	**2015 GC**	**2016 GC**	**2017 GC**	**2018 GC**	**Mean GC**	**SD GC**
Pure species controls							
AberBite (Lp4x)	7.0	6.5	5.0	6.5	4.7	$5.9^{b,c}$	1.0
AberMagic (Lp2x)	7.3	6.3	5.8	6.3	4.3	6.0^{c}	1.1
Caballo (Lm4x)	4.8	4.3	4.1	4.5	3.5	$4.2^{a,b,c}$	0.6
Podium (Lm2x)	4.6	4.5	3.8	4.1	3.0	$4.0^{a,b}$	0.8
Fure (Fp2x)	4.6	5.0	3.8	5.5	0.5	3.9^{a}	1.9
Kora (Fa6x)	6.2	3.1	5.0	5.8	3.7	$4.8^{a,b,c}$	1.3
LpFp							
Fabel	6.6	6.8	5.7	6.5	5.3	6.2	0.9
FuRs0142	6.3	6.5	5.5	6.1	5.0	5.9	1.0
Prior	5.5	5.8	4.5	5.3	3.5	4.9	1.0

Mean	6.1	6.3	5.2	6.0	4.6	5.7[c]	-
SD	0.7	1.0	0.8	0.8	1.2	-	1.1
LmFa							
FuRs0352	3.8	1.8	2.2	2.2	0.8	2.1	1.1
Bečva	4.6	3.5	3.6	4.1	3.3	3.9	0.7
Lofa	5.7	3.8	4.0	4.7	3.5	4.3	1.0
Mean	4.7	3.1	3.3	3.7	2.6	3.5[a]	-
SD	0.9	1.2	0.9	1.3	1.3	-	1.3
LmFp							
AberNiche	4.1	3.4	3.5	3.7	4.0	3.7	0.6
Agula	4.8	4.0	3.0	3.7	2.5	4.3	1.0
Felopa	4.3	3.8	3.8	3.8	1.8	3.6	1.1
Sulino	4.8	4.0	4.0	3.8	1.7	3.7	1.3
Hostyn	5.3	4.7	4.7	5.3	3.8	4.8	0.7
Perseus	5.5	4.7	5.0	5.7	4.8	5.1	0.7
Perun	4.7	5.0	4.3	5.0	3.5	4.5	0.7
Achilles	4.8	4.8	4.1	5.3	4.3	4.7	0.8
Mean	4.9	4.3	4.1	4.5	3.3	4.2[a]	-
SD	0.6	0.9	0.7	1.0	1.3	-	1.0
LmFg							
Lueur	4.3	2.7	2.3	3.2	2.1	2.9[a]	0.9

[a,b,c] Means within a column with different letters in superscript differ significantly ($p<0.05$)

Appendix 5. Pairwise comparisons for mean ground score cover between years

contrast	estimate	SE	df	t.ratio	p.value
2014 - 2015	0.85277778	0.16116935	41.9052586	5.29119081	3.94E-05
2014 - 2016	1.22013889	0.11323099	43.8176721	10.7756619	1.61E-12
2014 - 2017	0.45555556	0.12986536	51.3380887	3.50790655	0.00807482
2014 - 2018	2.24513889	0.16874102	66.7445902	13.3052348	0
2015 - 2016	0.36736111	0.1432625	43.9378863	2.56425165	0.09500125
2015 - 2017	-0.3972222	0.11311882	47.2348778	-3.5115486	0.00839385
2015 - 2018	1.39236111	0.14742419	69.9780075	9.44459073	0
2016 - 2017	-0.7645833	0.10408807	53.471652	-7.3455422	1.18E-08
2016 - 2018	1.025	0.1534562	63.353921	6.67943019	6.98E-08
2017 - 2018	1.78958333	0.12400305	56.9409938	14.4317689	1.10E-11

Appendix 6. Pairwise comparisons for mean ground score cover between cultivars of the same hybrid type for each year. Due to number of comparisons, only significantly different (<0.05) comparisons shown

year_1	var_1	year_2	var_2	estimate	SE	df	t.ratio	p.value

2014	Fabel	2018	Prior	3.1667	0.5096	150.1115	6.2144	0.0000
2015	Fabel	2016	Prior	2.3333	0.5096	150.1115	4.5791	0.0301
2015	Fabel	2018	Prior	3.3333	0.5096	150.1115	6.5415	0.0000
2017	Fabel	2018	Prior	3.0000	0.5096	150.1115	5.8874	0.0001
2014	FuRs0142	2018	Prior	2.8333	0.5096	150.1115	5.5603	0.0005
2015	FuRs0142	2018	Prior	3.0000	0.5096	150.1115	5.8874	0.0001
2017	FuRs0142	2018	Prior	2.6667	0.5096	150.1115	5.2332	0.0023
2015	Prior	2018	Prior	2.3333	0.4116	60.3205	5.6690	0.0015
2014	Bečva	2015	FuRs0352	2.8333	0.5096	150.1115	5.5603	0.0005
2014	Bečva	2016	FuRs0352	2.5000	0.5096	150.1115	4.9061	0.0088
2014	Bečva	2017	FuRs0352	2.5000	0.5096	150.1115	4.9061	0.0088
2014	Bečva	2018	FuRs0352	3.8333	0.5096	150.1115	7.5227	0.0000
2015	Bečva	2018	FuRs0352	2.6667	0.5096	150.1115	5.2332	0.0023
2016	Bečva	2018	FuRs0352	2.8333	0.5096	150.1115	5.5603	0.0005
2017	Bečva	2015	FuRs0352	2.3333	0.5096	150.1115	4.5791	0.0301
2017	Bečva	2018	FuRs0352	3.3333	0.5096	150.1115	6.5415	0.0000
2018	Bečva	2018	FuRs0352	2.5000	0.5096	150.1115	4.9061	0.0088
2018	Bečva	2014	Lofa	-2.3333	0.5096	150.1115	-4.5791	0.0301
2014	FuRs0352	2018	FuRs0352	3.0000	0.4875	48.0426	6.1537	0.0005
2015	FuRs0352	2014	Lofa	-3.8333	0.5096	150.1115	-7.5227	0.0000
2015	FuRs0352	2017	Lofa	-2.8333	0.5096	150.1115	-5.5603	0.0005
2016	FuRs0352	2014	Lofa	-3.5000	0.5096	150.1115	-6.8686	0.0000
2016	FuRs0352	2017	Lofa	-2.5000	0.5096	150.1115	-4.9061	0.0088
2017	FuRs0352	2014	Lofa	-3.5000	0.5096	150.1115	-6.8686	0.0000
2017	FuRs0352	2017	Lofa	-2.5000	0.5096	150.1115	-4.9061	0.0088
2018	FuRs0352	2014	Lofa	-4.8333	0.5096	150.1115	-9.4852	0.0000
2018	FuRs0352	2015	Lofa	-3.0000	0.5096	150.1115	-5.8874	0.0001
2018	FuRs0352	2016	Lofa	-3.1667	0.5096	150.1115	-6.2144	0.0000
2018	FuRs0352	2017	Lofa	-3.8333	0.5096	150.1115	-7.5227	0.0000
2018	FuRs0352	2018	Lofa	-2.6667	0.5096	150.1115	-5.2332	0.0023
2014	AberNiche	2018	Felopa	2.3333	0.5096	150.1115	4.5791	0.0301

2014	AberNiche	2018	Sulino	2.5000	0.5096	150.1115	4.9061	0.0088
2015	AberNiche	2017	Perseus	-2.3333	0.5096	150.1115	-4.5791	0.0301
2018	AberNiche	2018	Sulino	2.3333	0.5096	150.1115	4.5791	0.0301
2014	Achilles	2018	Agula	2.3333	0.5096	150.1115	4.5791	0.0301
2014	Achilles	2018	Felopa	3.0000	0.5096	150.1115	5.8874	0.0001
2014	Achilles	2018	Sulino	3.1667	0.5096	150.1115	6.2144	0.0000
2015	Achilles	2018	Agula	2.3333	0.5096	150.1115	4.5791	0.0301
2015	Achilles	2018	Felopa	3.0000	0.5096	150.1115	5.8874	0.0001
2015	Achilles	2018	Sulino	3.1667	0.5096	150.1115	6.2144	0.0000
2016	Achilles	2018	Felopa	2.3333	0.5096	150.1115	4.5791	0.0301
2016	Achilles	2018	Sulino	2.5000	0.5096	150.1115	4.9061	0.0088
2017	Achilles	2016	Agula	2.3333	0.5096	150.1115	4.5791	0.0301
2017	Achilles	2018	Agula	2.8333	0.5096	150.1115	5.5603	0.0005
2017	Achilles	2018	Felopa	3.5000	0.5096	150.1115	6.8686	0.0000
2017	Achilles	2018	Sulino	3.6667	0.5096	150.1115	7.1957	0.0000
2018	Achilles	2018	Felopa	2.5000	0.5096	150.1115	4.9061	0.0088
2018	Achilles	2018	Sulino	2.6667	0.5096	150.1115	5.2332	0.0023
2014	Agula	2018	Agula	2.3333	0.4875	48.0426	4.7862	0.0354
2014	Agula	2018	Felopa	3.0000	0.5096	150.1115	5.8874	0.0001
2014	Agula	2018	Sulino	3.1667	0.5096	150.1115	6.2144	0.0000
2015	Agula	2018	Sulino	2.3333	0.5096	150.1115	4.5791	0.0301
2016	Agula	2014	Hostyn	-2.3333	0.5096	150.1115	-4.5791	0.0301
2016	Agula	2017	Hostyn	-2.3333	0.5096	150.1115	-4.5791	0.0301
2016	Agula	2014	Perseus	-2.5000	0.5096	150.1115	-4.9061	0.0088
2016	Agula	2017	Perseus	-2.6667	0.5096	150.1115	-5.2332	0.0023
2018	Agula	2014	Felopa	-2.3333	0.5096	150.1115	-4.5791	0.0301
2018	Agula	2014	Hostyn	-2.8333	0.5096	150.1115	-5.5603	0.0005
2018	Agula	2017	Hostyn	-2.8333	0.5096	150.1115	-5.5603	0.0005
2018	Agula	2014	Perseus	-3.0000	0.5096	150.1115	-5.8874	0.0001
2018	Agula	2016	Perseus	-2.5000	0.5096	150.1115	-4.9061	0.0088
2018	Agula	2017	Perseus	-3.1667	0.5096	150.1115	-6.2144	0.0000
2018	Agula	2018	Perseus	-2.3333	0.5096	150.1115	-4.5791	0.0301
2018	Agula	2015	Perun	-2.5000	0.5096	150.1115	-4.9061	0.0088
2018	Agula	2017	Perun	-2.5000	0.5096	150.1115	-4.9061	0.0088
2018	Agula	2014	Sulino	-2.3333	0.5096	150.1115	-4.5791	0.0301
2014	Felopa	2018	Felopa	3.0000	0.4875	48.0426	6.1537	0.0005
2015	Felopa	2018	Felopa	2.0000	0.4116	60.3205	4.8592	0.0227
2017	Felopa	2018	Felopa	2.0000	0.4023	41.1632	4.9717	0.0255
2014	Felopa	2018	Sulino	3.1667	0.5096	150.1115	6.2144	0.0000
2018	Felopa	2014	Hostyn	-3.5000	0.5096	150.1115	-6.8686	0.0000
2018	Felopa	2015	Hostyn	-2.8333	0.5096	150.1115	-5.5603	0.0005

2018	Felopa	2016	Hostyn	-2.8333	0.5096	150.1115	-5.5603	0.0005
2018	Felopa	2017	Hostyn	-3.5000	0.5096	150.1115	-6.8686	0.0000
2018	Felopa	2014	Perseus	-3.6667	0.5096	150.1115	-7.1957	0.0000
2018	Felopa	2015	Perseus	-2.8333	0.5096	150.1115	-5.5603	0.0005
2018	Felopa	2016	Perseus	-3.1667	0.5096	150.1115	-6.2144	0.0000
2018	Felopa	2017	Perseus	-3.8333	0.5096	150.1115	-7.5227	0.0000
2018	Felopa	2018	Perseus	-3.0000	0.5096	150.1115	-5.8874	0.0001
2018	Felopa	2014	Perun	-2.8333	0.5096	150.1115	-5.5603	0.0005
2018	Felopa	2015	Perun	-3.1667	0.5096	150.1115	-6.2144	0.0000
2018	Felopa	2016	Perun	-2.5000	0.5096	150.1115	-4.9061	0.0088
2018	Felopa	2017	Perun	-3.1667	0.5096	150.1115	-6.2144	0.0000
2018	Felopa	2014	Sulino	-3.0000	0.5096	150.1115	-5.8874	0.0001
2014	Hostyn	2018	Sulino	3.6667	0.5096	150.1115	7.1957	0.0000
2015	Hostyn	2018	Sulino	3.0000	0.5096	150.1115	5.8874	0.0001
2016	Hostyn	2018	Sulino	3.0000	0.5096	150.1115	5.8874	0.0001
2017	Hostyn	2018	Sulino	3.6667	0.5096	150.1115	7.1957	0.0000
2014	Lueur	2016	Lueur	2.0000	0.4000	37.8596	5.0000	0.0264

Appendix 7. Pairwise comparisons for mean yield between years

contrast	estimate	SE	df	t.ratio	p.value
2014 - 2015	-3.4332144	0.18918192	46.2463517	-18.147688	0
2014 - 2016	0.51681037	0.20120762	21.4016891	2.56854275	0.11241351
2014 - 2017	-1.6443402	0.23079725	44.8882221	-7.1246092	6.64E-08
2014 - 2018	1.02235219	0.24966421	23.7566568	4.09490888	0.00351045
2015 - 2016	3.95002475	0.18390082	50.6120289	21.4791037	0
2015 - 2017	1.7888742	0.1871537	57.731573	9.55831593	1.57E-11
2015 - 2018	4.45556658	0.27702434	75.2798891	16.0836645	0
2016 - 2017	-2.1611506	0.22230151	51.3615298	-9.7217088	0
2016 - 2018	0.50554183	0.23921519	36.2726765	2.11333495	0.23653939
2017 - 2018	2.66669238	0.26824327	64.414876	9.94132073	6.64E-12

Appendix 8. Description of F2 mapping family, sequenced using the GenoGrass approach, however not further analysed as part of this study due to the few number of variants generated to perform mapping. The description contains a description of the populations, the associated publications as well as the number of variants which were present in the final SNP set post filtering

Group	**Sample number**	**Sample type**	**Publication**	**Population type**	**RD**	**Variants**
F2	101	Individuals	Anhalt *et al.,* 2008 Anhalt *et al*., 2009	F2	Mean=131.59 Min= 0	

			Tomaszewski *et al.,* 2012 Velmurugan *et al* ., 2016		Max= 6962.18 IQR= 302.08	

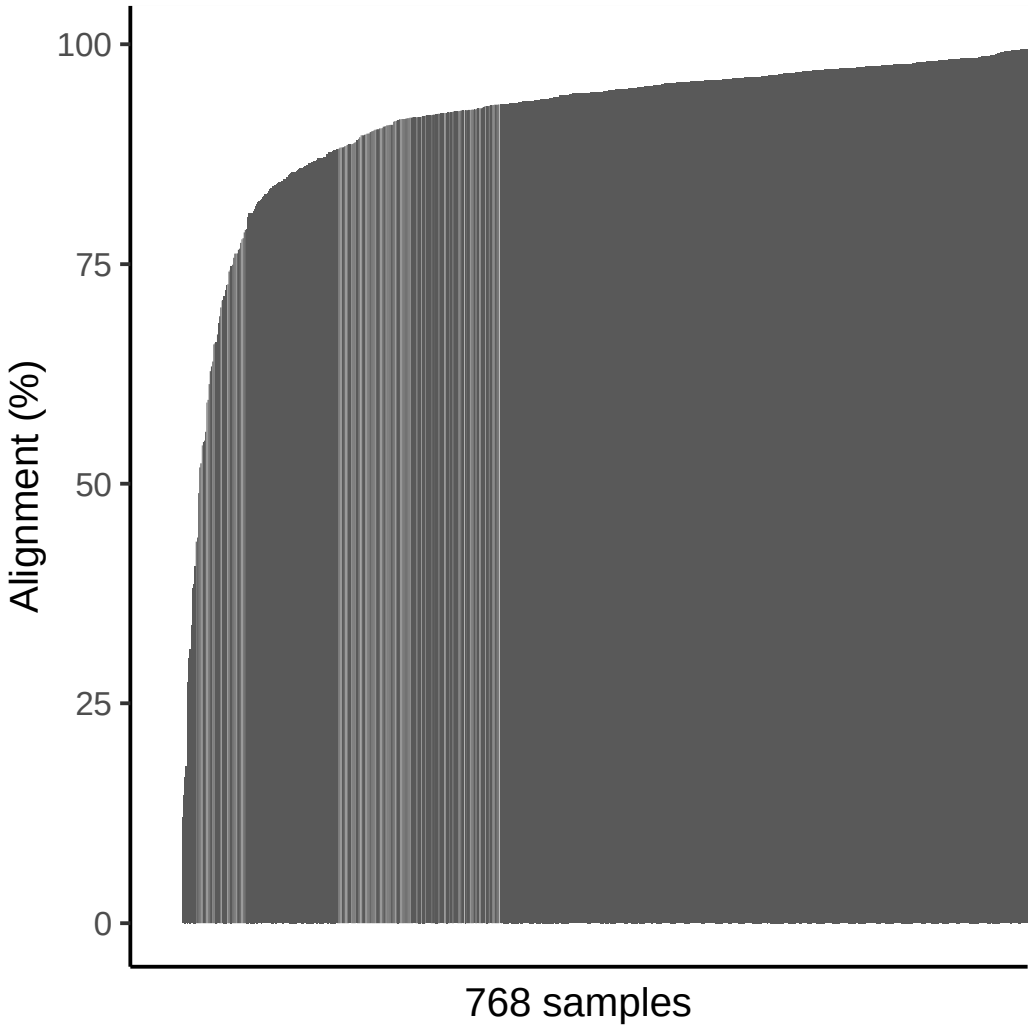
Appendix 8. Targets included in the array, the scaffold the and the their associated primer

Scaffold	Start	End	Forward primer	Reverse primer
scaffold_1|ref0040986	42394	42603	AGTCGCCGAAGCTGAAGAC	CCACTTGGGACCTGATTCTG
scaffold_100|ref0001926	68915	69093	GGTACACCAAGACGGCAAAG	TTCCCCTGCACCCTTATTG
scaffold_102|ref0001924	197965	198158	TCATCCTCTTTTGGGCTGTC	TAGCAAAGCGGTAGGGAGTG
scaffold_10240|ref0008205	24861	25067	TCCCACGACATCCTCTTGTC	GATATGCAGCGACGTGAAAG
scaffold_10324|ref0009708	4932	5102	GCCGCCTCTTAGTCAGCAC	TTAGCTGCCTTCTTGGCATC
scaffold_10899|ref0029618	2337	2510	AGGCGCTAAGGAGGATGTTC	CCACTGCTCACTTGTTGGTC
scaffold_10942|ref0037446	31838	32033	AACTGTGGCCTGGTATGGTG	GTACTCCGCCGTCTTCACC
scaffold_110|ref0032459	3390	3542	GCGTCGGTCTTTGTGTTCTC	CATTACCAGCACCACACAGG
scaffold_11281|ref0023388	24073	24272	GTGAACAGCCTCCCACCTC	GCAGCAACTCCACAAGCAC
scaffold_11333|ref0028119	5381	5562	CGAGGGCATGAGGTAGAAGC	GCCATTGCATCCAGCTTAC
scaffold_11466|ref0003439	16872	17074	GCAGCTACTAACGGTGCTTTC	CACACAAGTTCATCAGTGCAAAC
scaffold_1149|ref0040345	34743	34944	TCTTGATGCCGGAGATGTTC	GTTGCAGCCACCTTTGATTC
scaffold_1153|ref0009992	109916	110102	ATGCAGGTAGATCGAGACAGG	CTCTGCATTCACCGATCAGC
scaffold_1161|ref0004474	133565	133760	CGCAGCAGATAAAAAGGAATG	AAGCATACAACAGCCCAACTG
scaffold_119|ref0032466	202010	202174	ATCCATCCATCCATGTTTCG	TTGCGGGCTGTGAAATAAAG
scaffold_11945|ref0043388	4510	4680	GATGGTGGTGAAGCAGATCC	CAGCAGAAAAACGAGCAAGC
scaffold_11991|ref0000373	27909	28098	CTCGAAACCCGAAGAAACTG	TCGAGCAGCCACATCTAAAG
scaffold_12079|ref0011165	15560	15767	CCTTCCTGTGTCCCGAGAG	CGCAGATCGTTTTGGTATCC
scaffold_12193|ref0024469	24930	25142	GGACGACGACGGCTTCTTG	GGAGAAGCCGAAGCTGGTC
scaffold_12432|ref0025169	17938	18135	CAATCGAAACCCAGTCCATC	CAACCTGTCGGACTTCTTCC
scaffold_12504|ref0045352	8048	8241	ACAAGAAGGTTCCGTGGATG	TTCGCCGATCTTATTTCCAG
scaffold_12531|ref0002237	23281	23431	ACTTCTTGCGTCCTCCTTCC	AAAGTCGCAACACCCACAAC
scaffold_1293|ref0004481	9000	9168	CTGCCACATGAGACCGAAC	TGCTAGTGTTGAAGCGTTGC
scaffold_13063|ref0011328	89716	89909	GCGGCTCTGTCTCACCATATAC	CCTGCTAAGAACTGGGAGGAC
scaffold_1324|ref0023232	11550	11727	TGGCAAAGTGAGATGAAAAGTG	GTGTTTCAGCAGCAATGAATG
scaffold_13308|ref0034440	12273	12428	CCTGCTAGATGGTTTTGATGG	GTCTGAAGCCAACAAGTCTACG
scaffold_13366|ref0021945	816	989	TAAGCCCCAGGCTCGTATC	ATCGGTGATCGCTCCTAGAC
scaffold_1343|ref0010619	10812	10979	GCAACCTAGCGAAAAGGTTTG	ATCGTTATCCCGTTCTGCTC
scaffold_13475|ref0026757	53730	53923	GGTGTTGTGGAGGTGATCG	CCTGAAACAGGCAGGAGTTC
scaffold_1368|ref0046522	6124	6326	TGGACCTTCCTTTCAGCTTG	AGTATTAACGAGTTGAGTTGAGATGG
scaffold_13819|ref0008891	54344	54528	AAGCCACTCCAGTTCGTTTC	AGCATGACTTATCGGTTGGAC
scaffold_1422|ref0033620	89675	89843	TGAAGTTGCAGAGCGTCTTG	CAACAGCAACCAACTCAACG
scaffold_14281|ref0021019	10208	10416	CTGAAACCTGATCGTGATGG	GGCTCAAGCATCCACATTAAC
scaffold_14299|ref0018060	32973	33169	TGAACTAGCAAAATTAGTATGCGAAC	CTTGAAGGCTGCTGAAGAATATC
scaffold_14962|ref0035979	20132	20325	CTCGCAACAAGAGTGTGAAAC	GCTGCTTTCCCTTTCTTTACC
scaffold_15042|ref0029108	11898	12101	AGATCAAGCTGCCCAAGGAC	CCAGGAAGATGTACGTTGGTG
scaffold_15053|ref0047046	10774	10960	CTCTTGGAACGGGATGATTG	GTAATGTTCATCCCGCACTG
scaffold_15123|ref0001580	9686	9899	GTCTGAGCTTTGCTCGGTTC	CTCAAATCCAAGCAGGTTCC
scaffold_1519|ref0023932	12319	12495	AAATCGCTGGTGGGTTCTC	GACCACGGAGATCCTCTTCC
scaffold_15267|ref0023130	1768	1972	CTCCTATCGTTCATCCTTTTGC	TGGCTAGATTGTTCTGGAAGC
scaffold_1565|ref0043379	10604	10811	CAATCAACGGTCAAATGCAG	TAGGTCGGCAAGTCTGTTGG
scaffold_15670|ref0043540	119748	119933	CATGGTCGATTCCAGACCAC	CAACCATCCGCAGCAGTAG
scaffold_16787|ref0028195	6819	7023	GCGAAACAGTGCTCGATAGC	GGCAGCAAGAGCAAATATGAG
scaffold_16870|ref0006007	10360	10549	AGAGGTGGCTTAGGGTAATGG	CACGAAGAAGCTCAACATGC
scaffold_1712|ref0000672	62446	62604	CAATGTGCCTGCCAGTTATC	CATGCTGGAAGTGGCAGTAG
scaffold_17310|ref0004132	6239	6408	ACCTGGAAGAGGCTGGTGAG	CGTGCTCGACTACCTCTACAGC

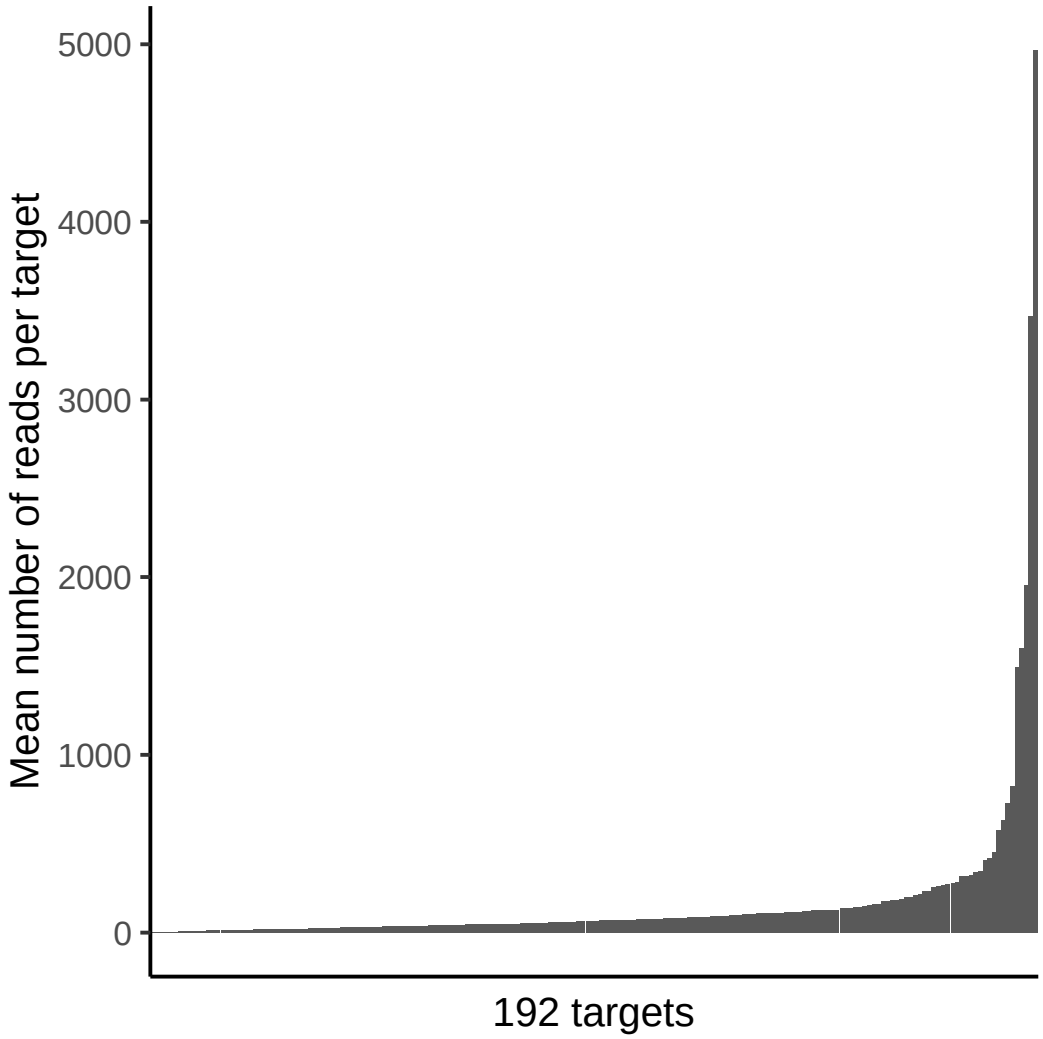
scaffold_176|ref0043339	98912	99103	CTAGCCTTGCCTCCTGATTC	TCATCTCATCCTCTGTTTCTGC
scaffold_17729|ref0038942	6257	6412	CTGCATCGTCCTCCTCCTAC	CGCACTCTCCTGTTTCAGTG
scaffold_1784|ref0029502	7413	7594	GATGCTTCTCTGAGGTGGATG	TTTCCACTCAAAGCACTAGCC
scaffold_17937|ref0016356	49684	49872	CATCGACGAGAATGGAGCAG	CCATCACAGCAGCAAGAGAG
scaffold_18|ref0042669	12462	12624	TCCGTAGCTGTTTTTGGAATG	TGTGCAATTCTTTTCCTGCTG
scaffold_182|ref0000331	60891	61073	AGATCAACCGGGAGGTGAC	CGCCCTTCTTCCTCCAGTC
scaffold_1865|ref0047453	1945	2142	GAAAGGGCGCTCACATAGC	TCTTCCCCTGGCACTCCTG
scaffold_1872|ref0029546	32743	32919	ATTTGGAATCTTGAGAAAAGAAGG	AAGCAGCAGTTTTTGACCTTTC
scaffold_18961|ref0022259	5639	5844	CAGGCTGCAAAGAGATTCG	AGGTGCAGGACTTCAACAGC
scaffold_19449|ref0033434	49365	49566	CTGCTCACATCGGATTCAAC	TCATCAAGCATCATCAGCAAG
scaffold_1959|ref0042697	26214	26394	CATGTATGTACGGCTGCTTCC	CCAGCAACAGACCAGTTCAG
scaffold_1961|ref0037400	67925	68091	CACAACAAAAGCGAACATGC	ACTCCGATTTTGACCACGAC
scaffold_1989|ref0003138	54326	54547	GAGGGATTGGGTGAGGTAGG	CTAGCCAACCCTATGCGATG
scaffold_200|ref0035631	1958	2157	CTGAGGATAGCAGCCTGGAC	ATGGAGCCAGTGGAGAAGTG
scaffold_2020|ref0025718	718	893	CCACGGGCTAATTTACAACC	GGCTGTTGCTGCTGTTGTC
scaffold_2049|ref0038154	51608	51813	GTTGCCGACCTCTCCTTCC	CGCACGGTTCAAGATGTCC
scaffold_2075|ref0020174	96268	96441	TGTTCGTGACCCACAGAGAG	GTCGTGGTCCTGAATTGGAG
scaffold_2084|ref0014659	72262	72471	CTCCTCCTCGCCACCTTC	GAAGGAGCGGTGGGAGAG
scaffold_2093|ref0032591	216982	217169	CCAACCATTGGGTCAGTTTG	CGAAACGAATTGGAGACAGG
scaffold_2099|ref0032593	113094	113294	TCCACCGCCACTCCAAAG	CTCTCGGAGGATCAATACCG
scaffold_2110|ref0018532	39061	39244	GCAGACTGATTTATGGGCTCAG	TGCTAACAGAAGGCATAATGAAAC
scaffold_2145|ref0024453	17338	17506	GGCCGTTGTTGTTCTGAGTC	TGTTGTCATCAGAGGCATCG
scaffold_2166|ref0015754	73359	73525	TTCACCGCTACACCAGATTG	AATTCAGTTCCCAGCCACAG
scaffold_2191|ref0033664	30113	30305	CGGAACAAGTAACGTGGAAAG	TGTCGCTAAACGAACACCAC
scaffold_220|ref0023189	84768	84959	TGCTTCTTCAGCTCTCATGC	GATCGCAACAGGAAAGCTG
scaffold_2264|ref0002514	9551	9761	CTCCGCAGCAAATTGTCAC	ATACTGCTGAGGCGCATTTG
scaffold_2268|ref0002502	6264	6451	TGCATCTTCACACCTCAACC	CCGGGCACTTCTTTTGGTAG
scaffold_23054|ref0020106	109573	109752	GGGCGACATGCTTATCCTC	GGCCGTGGAGTGGTAGAGG
scaffold_2332|ref0018206	99199	99359	GTGAGGTGGTCCTCGACATC	AAGTAGAGGGCGTGGTTCG
scaffold_2344|ref0000209	56673	56837	CATCGCAGAACATGAAGCAG	TCTTTGACCCAGTGCAGGAC
scaffold_237|ref0040978	162610	162774	GACAAAGCGAGCGGATCAC	ACAAAACAATGCCCTCATGG
scaffold_241|ref0012154	12692	12861	CTGAGAGGCTAAGGGAGACG	ATGCTCAACCACTCCTGACG
scaffold_25347|ref0006749	13596	13782	TGGTTGGTGGGGATTTAATG	TTGACTGGAGCAGAGGGATAG
scaffold_2543|ref0034320	1454	1629	CTTGGATGCTCGTTGATGG	CAAGGAGCAGCACATGAAAG
scaffold_25549|ref0039256	75124	75329	ACGACTTCGGCACGATCAG	GGAGGTGTTCGAGGACGAG
scaffold_2638|ref0048145	76795	76988	CAAGAATGGCCGCATCATC	CACCGGCCATCATAACCTC
scaffold_2694|ref0013918	60075	60259	CTTCTAGGGGCTCGTCACAG	ACTTCCTGTGCCACCATCAC
scaffold_2702|ref0027944	56474	56654	ACTGCCATCAACATCAATTCAC	GCAGCAGATGAAGCACGAC
scaffold_27686|ref0045623	62313	62513	GTGTTCAGGATTCCCCAGAG	AACGGTCCAAGTTCTGGTTG
scaffold_2778|ref0045866	43296	43492	ACGCTGTAGTGTGCCGATG	ATTGCAGCCCGTCTCCTC
scaffold_2801|ref0025790	3583	3775	CTCCTCGAACTACGGTGCTG	AACCCCGCTCCTACAAATAAG
scaffold_2846|ref0029400	48366	48576	TAGCCCCTACCTTTCCATCC	GCAGTCTGTGGCTGCTCTC
scaffold_29285|ref0008041	2674	2861	TGGCTATCAGATCAGGACGAC	GCCTGCCTGGGATAATCTG
scaffold_2949|ref0012763	64565	64723	CGCAAACCATCTTCTCACAG	CTAACGGCCAGTGCTTCTG
scaffold_2959|ref0043237	37944	38117	CTCTCGCTCGGCTTCAATG	AGCCCTTTGGTGCTGCTG
scaffold_30012|ref0027143	102732	102902	TGCCGAATCTGCCTTAGATG	GATCATTGACGGGATTTTCC
scaffold_3038|ref0025552	1465	1652	AACATCTCTAACGGCCCTAGC	ATATTAGCGCACGCAGTCAG
scaffold_3049|ref0007382	58379	58551	GCTAGGAGTGGTCCGAACTG	GCGCTCGACCTGGAGATG
scaffold_3061|ref0020001	7284	7496	TAGTTTTGTGGCCCTTCTCC	GCACCACTGTCACTGTCACC
scaffold_322|ref0033181	78966	79165	TTTGGGCATAAAGTTCTCGTC	AGTGGACACCCCCATGTTAG
scaffold_3232|ref0030065	1919	2101	CATCTGGGGAGAACTGAACC	CTCCTTTTGTTTGGCCTTCC
scaffold_3304|ref0020202	24919	25131	CATACCGCCGAAATTAACACC	CCGCGACTAATCCTCTCCTC
scaffold_3350|ref0014712	61663	61835	GAGTGAAATGATGAGGCAAGG	CATAGGCACGTTGTTCTCAGG
scaffold_34|ref0030224	996	1190	CTCGCTGGATTTCGTCAAC	CAGCCATCATTACTTCGATGC
scaffold_342|ref0045820	49658	49868	AGCTTCCAGAACACCGAGAG	CTGGAAGGTTTCTCGTACCG
scaffold_3598|ref0021228	30366	30543	ATTGTCCTCCGGCTCCATC	ACGTAGGGTAGGGCGTTGTC
scaffold_360|ref0009959	19329	19549	AGAACTCGCTGGGGAACAG	CATTCTCCCCCTGCACTTG

scaffold_36762|ref0028971	103127	103338	TCACAAGGACGGCTCTCAG	TACATCCCAGATGGCTCAGG
scaffold_3732|ref0013341	66721	66889	GGGAGAAGGGTGTCAGAGTG	TCTCAGCCATCTCCCTGTG
scaffold_384|ref0019067	91985	92147	TGGTCAGCAAAGGAAGAGAAG	CCTGATGTTTGTCCGTATTGTG
scaffold_413|ref0039281	54110	54283	GCTGCAACAAGAGTGAGCTTC	GCGGAGGTTCTAAAAAGTGAGG
scaffold_4166|ref0045065	23231	23441	CATACACAGTTGGTCCCTACACAG	TTTTGACAACATCCGTCAGG
scaffold_4234|ref0021515	73686	73869	TGCAAATTAGACGCAGCTTTG	ATGATGTGGTCTGCCCTAGC
scaffold_4369|ref0002961	466	624	AGATTCAGCAGGAGCAATCTG	CAGAGCAGCAGGAGAGAGTG
scaffold_4448|ref0017406	51228	51405	CCAAAGCCCAAAAATGTAACC	CGATCTGCTACTCGTGGTACG
scaffold_4576|ref0045556	40	216	GCAGCGTTAGCACTTAGCAG	GACGCTTTTACCCAAGATGG
scaffold_4630|ref0024910	6380	6581	CAGTAGACCACGCCTGATATTTG	GATGGATGGGATCAAGCTG
scaffold_464|ref0008782	87187	87364	GCTTGGTAGAGGTGGTCTGC	GCACCTCGTCTCATTCACG
scaffold_4659|ref0037345	129952	130130	TCATCGTCAGGGAAAGGATG	TTAGCCGCAGCCATCATATC
scaffold_47|ref0001307	16839	17022	TGACGCGAGAGATACCACAG	GGTTCCAGTTTCGCTCCTTC
scaffold_4717|ref0037798	27714	27911	ACAAGCAGCAGCATCTCCTC	AACAATCGCACTAGCAGCAG
scaffold_4786|ref0005582	4868	5047	GAGGCAAGACAAAACGGAAC	AGTTGCGTGTCGAACTTGTG
scaffold_4902|ref0021609	51630	51789	CTTGTGCCACCCCAGATTG	TAGAGGCGCGGTTTGTTAAG
scaffold_504|ref0006340	2144	2337	AATGTGACCCCTGCTGTTTC	GCAGAAAGGAAGGGGATGTC
scaffold_51|ref0019419	20720	20922	CATACAATTTCCCCCAGGTG	ATCCGTCAAGGCTGACCTTC
scaffold_5122|ref0013885	51126	51329	CGAAGGTCCGAGCAAAGTAG	CATTCCGGCTCAGCTATCTC
scaffold_5148|ref0026632	14884	15085	AGCTCTTGCCTCATTTTTGG	AGAGATTCCCTCGTGGTGTG
scaffold_5229|ref0045651	161095	161304	CGAAGGGCAAGAGTGCTAAG	CTTCGACGATGCAGCTTTC
scaffold_527|ref0042154	40047	40251	TTCATCGCGGGCTTTTATAC	ATAACCAGCCCACCTAGCAG
scaffold_5301|ref0047902	40479	40670	AGGTCTTTTGTAGGTGGAAACG	CTGGTCACTGCTGCTCTCTG
scaffold_5346|ref0024732	22543	22744	TAAATATCAGCGGGGAGAGC	ATATGGTGTGGCGTGCAGAG
scaffold_5369|ref0011976	11919	12089	ATAAGTCGAACTCCGCAACG	TGGAGACGCTGAAGACAGAC
scaffold_5403|ref0046155	9478	9675	CAGTCTACGTTGGCGGAATC	ATCATGGGGAACAAGACGAC
scaffold_5599|ref0028885	27898	28091	TTGTTTGGATGGGATCAGTG	GGAGGAACTGCGTCTACGTC
scaffold_6043|ref0000269	44081	44256	CAGGACGCAGCCAAGTCTAC	CAGAAGCAAACTGTGGATGAAC
scaffold_6149|ref0035523	10045	10227	CTCTGAACCGAAACATGCAG	CTGGAAGGGATGTTCTTTCG
scaffold_615|ref0009431	42879	43053	ACAGGTGGTAGGTGGCAAAG	ACGAGGCGGAACAAAAATG
scaffold_6177|ref0005114	23541	23725	GACAATGCACAAAAACCTGAAC	TATGCTCAGGCTGCTCAAAG
scaffold_6246|ref0025895	60813	61009	TGTGCCTGTGGAAAAAGGTC	GGCTTCAGCTTCTCACTTGG
scaffold_6382|ref0000970	55079	55250	TCCAGCCCACATACTGTTCTC	TAAGTGAAACTGTCCGTAGAATGG
scaffold_6436|ref0002950	46496	46672	GCGGTAGCAGAGAGATGGAC	CGACTGAAGCAGCGTCTTTAC
scaffold_650|ref0032730	20902	21070	TCCTCTCGCCGTACTTGATG	TCCTGTTGCTTCACCTTGG
scaffold_6578|ref0047298	58674	58837	ATGAAGAGCGCCAAGAGAAG	CTTCCGCCTTGGTCATGTG
scaffold_6609|ref0039591	25904	26104	CTCGTCGTTTCCTACGGTTC	AAGAAGCAGCACCCAAACAG
scaffold_6759|ref0031818	97520	97714	ACGCCACCATTCCAGTGAG	GGTAATCGCTGCAAACGTG
scaffold_6767|ref0013818	14957	15127	CCATGCGATGATCCACTTC	CTTGGATGATGGGAGAGTCG
scaffold_6784|ref0026550	45742	45916	ATCTATCCTCGCCCTGATCG	TGCACAAATGGTTATCACACC
scaffold_6798|ref0008476	33850	34029	CACCTGGTCCCTGTACTGATTC	GCACGGCAAGATGAATACTG
scaffold_6820|ref0032280	42496	42684	GAGAACCTCGAAAGCAGCAG	GACTGCTGGACTGCCAAGAC
scaffold_6906|ref0004902	86757	86958	GCAGGTTGTTGGACAGGTTC	CTGATGAACAACTTCCTGTTCG
scaffold_699|ref0011199	33755	33930	GAAGCCTGTTTTTCCACCAG	CATTTCGCTGCTTCCAAAG
scaffold_7178|ref0038347	43151	43322	TGGGACAGAAATGAATCGTG	ACATCGAAACCACGACCTTC
scaffold_7257|ref0024869	12767	12974	TCACGTTGTTTCCCATGTTG	CAGGCATGGAGAGGAATCAC
scaffold_7312|ref0001899	8501	8698	CAAGACGGGGTCGTGGAG	GTCCTCTTCGGGGGTGAAG
scaffold_7328|ref0045034	22757	22955	GGCTGCAAAGGTAAGAATACG	GATCCTCATGGTGTATTTGTGC
scaffold_7361|ref0020044	7071	7275	ACACCAACAATCCAGCATCC	TTATTGCAGGTCAGCGTGAG
scaffold_745|ref0048287	6281	6453	ATGATGAGGACGAGGGAGTC	TGCGACCTCACCTACTCCTC
scaffold_7641|ref0036653	11989	12209	GGATAGTTATTCCGAACTAGAGCAC	CGAATGCACCACTCTGAACC
scaffold_776|ref0042890	127840	128059	GCTAGTTGTGGTGGCATGG	ATGCTGCAACGTGGTCCTC
scaffold_7767|ref0017329	35734	35920	TTCCCTGCGAAATATCCATC	TAGGTTCGCTCAGGAGTTCG
scaffold_780|ref0026871	34847	35012	GATTCGCCCTCGTCTCTATG	CTTCAGAGGACTCGGTGAGC
scaffold_7903|ref0045601	27137	27309	CGCATAGCTTGGTGAGCTG	GGAGTGAGGAAATAGAGATGCAG
scaffold_7960|ref0007977	31719	31893	AAGGTCATCGGAGGAGGTG	ATTTAAGCACAACGCCATCC
scaffold_8121|ref0015158	20092	20302	TCATCGTCGTCATGTCCATC	GCTTCTGATTCCGACGAGAC

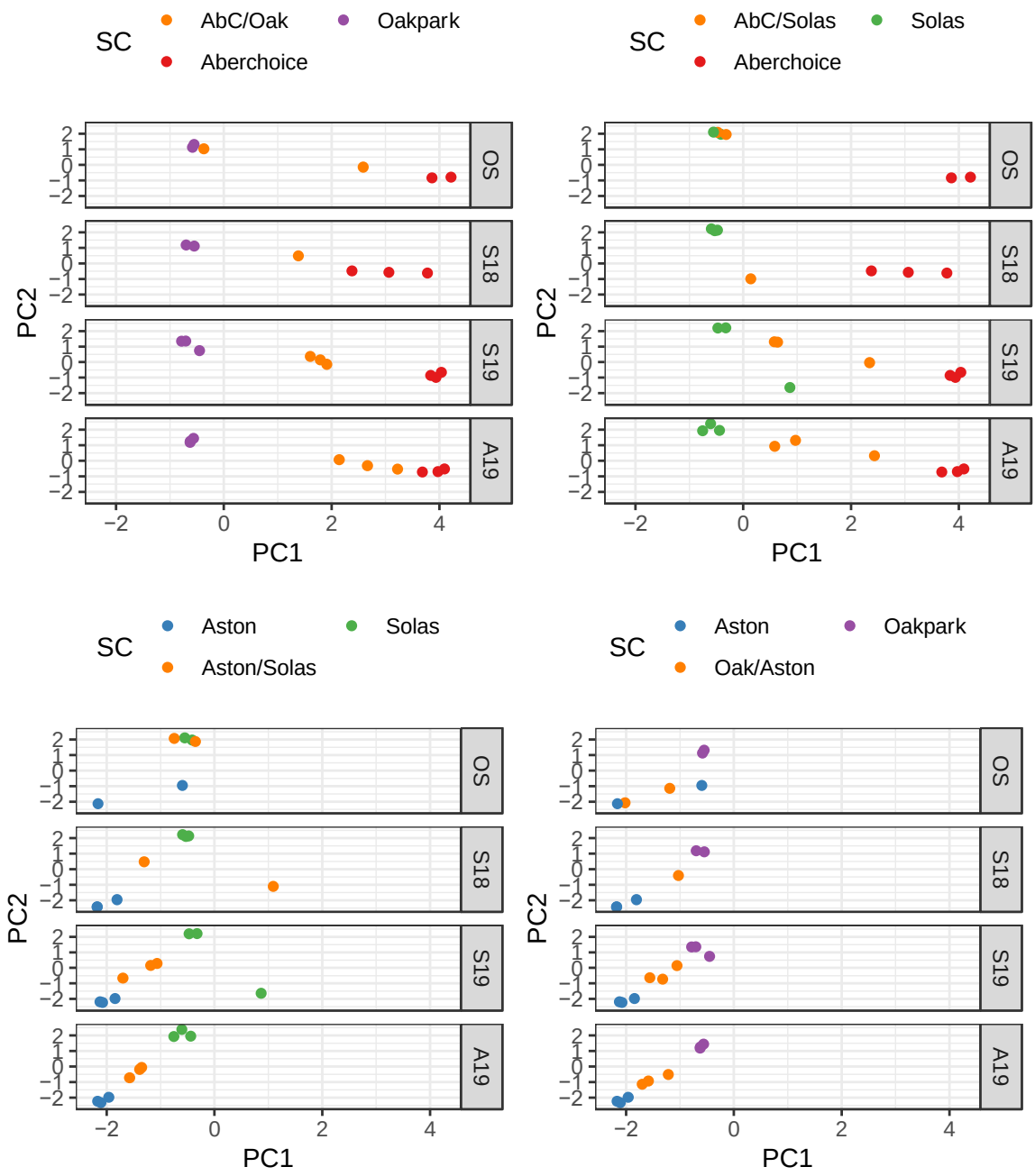
scaffold_8189|ref0000840	12481	12638	GACGGGTGGAAGAGGTAGC	CGGACCAGCTCTTCTACTGC
scaffold_8192|ref0031454	40329	40527	GTCCTTGCGAGTAGGAGCAG	TCTAAGCGTCGGACCAGAAC
scaffold_8249|ref0017106	25825	25995	CATGTATCCCTGGAAGACTGC	TTGTGGCTGAAAACTGCAAG
scaffold_8261|ref0029690	11512	11714	TCATACAATTCTGGGTCTCACG	GTTGCGCTTTGCTGAATATG
scaffold_8334|ref0001886	166193	166372	GCACATAGTCCAGCCCATC	GCATCAAGCCATCAAGTACATC
scaffold_8529|ref0019588	30266	30476	ACCGTCCAGGTCCAGCTC	GGTGGGGTCTGCTTCACC
scaffold_8594|ref0039194	21548	21755	GGTACTCCAAGCCAGTGAGG	ACAAGCTCGCTGGAATCATC
scaffold_8646|ref0026746	15664	15856	TGCACCAGAGTAACAAAGATGAAC	CCTTCCTATGGTGGGATTTG
scaffold_87|ref0023176	14657	14831	CTCTTCGGGACTTCGCTTC	CTGCCCACACATCCCTAGC
scaffold_8726|ref0024402	50450	50641	TGGCGCAAATTCAGAAGAAG	TCAACATCTACAACAAGCAGGTC
scaffold_8917|ref0017651	14015	14174	TCGTGTGCTCTAGGGTTTCTC	GGATCACAAAGAGGCCAATC
scaffold_8984|ref0037328	37969	38149	CAGGGTTCATCTCCGAACAG	CGCCTAATTCAAGCAGCAG
scaffold_9|ref0023150	37146	37350	CTGCTCCCCCATGAACAG	GTGGGGACCAGCAATCTG
scaffold_908|ref0007599	8371	8558	AGGTGAGCAGCCTGAACTTG	CTGGGACTAACAAGTGATATTTGC
scaffold_915|ref0038168	13247	13432	GAAGCGTGTGACGATTGATG	CAGAACGCGCCAGTGAAG
scaffold_917|ref0038170	21330	21515	GTGATTGCATCCGAGGTAGC	GATCTTGCTCGTGCTCTTGC
scaffold_9199|ref0014555	185209	185406	CTGGCACTGTCACCAAGATG	AAAGAGATGGCTGGGGAGAG
scaffold_9278|ref0023533	16714	16900	ATTCAGGGGCATCATTGTTC	ATACCTCCGGCCTAGAATGC
scaffold_9309|ref0020396	3616	3788	GCTGCCCAGATAACACTTCC	CCTGCGAGTGTTATGGTTTG
scaffold_9446|ref0047393	9452	9620	ATTTTCTTGCGGAATTGTCG	ATTTCTGGGAGGTGGGAGTC
scaffold_9561|ref0028232	338922	339112	CTAACAATCATCCCGGAAGC	CCTGGTAGCTGGTCTTCTGG
scaffold_957|ref0012939	110881	111075	GTGCCCCATTGTTTCTCTCC	GCAGCTCCAATCAGTTTTCAC
scaffold_980|ref0023409	39220	39421	AACAGACGCAAACCGTAAATG	TTGAAACCCCTCGAAGACG
scaffold_989|ref0007130	106672	106864	GAAGAGGAGACCGCGACAC	TTGGTGAGCTGCTACTCTTCC
scaffold_12573|ref0027446	34808	34989	CTGGGTGATGAAGCTGTGAG	CAGCAAACAGCAACGGTATG
scaffold_1374|ref0015999	28930	29110	TCCTTCTCCAGACCATAGGC	CAGCAGACCCTCGATGAGC
scaffold_14822|ref0021055	30710	30883	TCAGAAATCGACACCATTGC	TTGTTCTGGTACGTGGGTTG
scaffold_15522|ref0012816	39423	39594	GGTCACCGTCTGCTGTCAAC	GCTGTTGGCCTGATGCTAAG
scaffold_1627|ref0028949	37584	37779	GACATCTGCCACTTCTTCACC	GTCCTGCTGAGCTGCTACG
scaffold_2183|ref0003185	85486	85677	GACTTCAGCAGTGGGAGAGG	AACTAACCGGCAACTTGTGC
scaffold_2318|ref0005675	131474	131659	ACTCCCTCCAGGTCACCATC	CTCGGTGGAGAAGATGTTGG
scaffold_250|ref0030177	107589	107752	GATAGCAGCTCCCTGTGTCC	TGATTCAGAGCGAGCAGATG



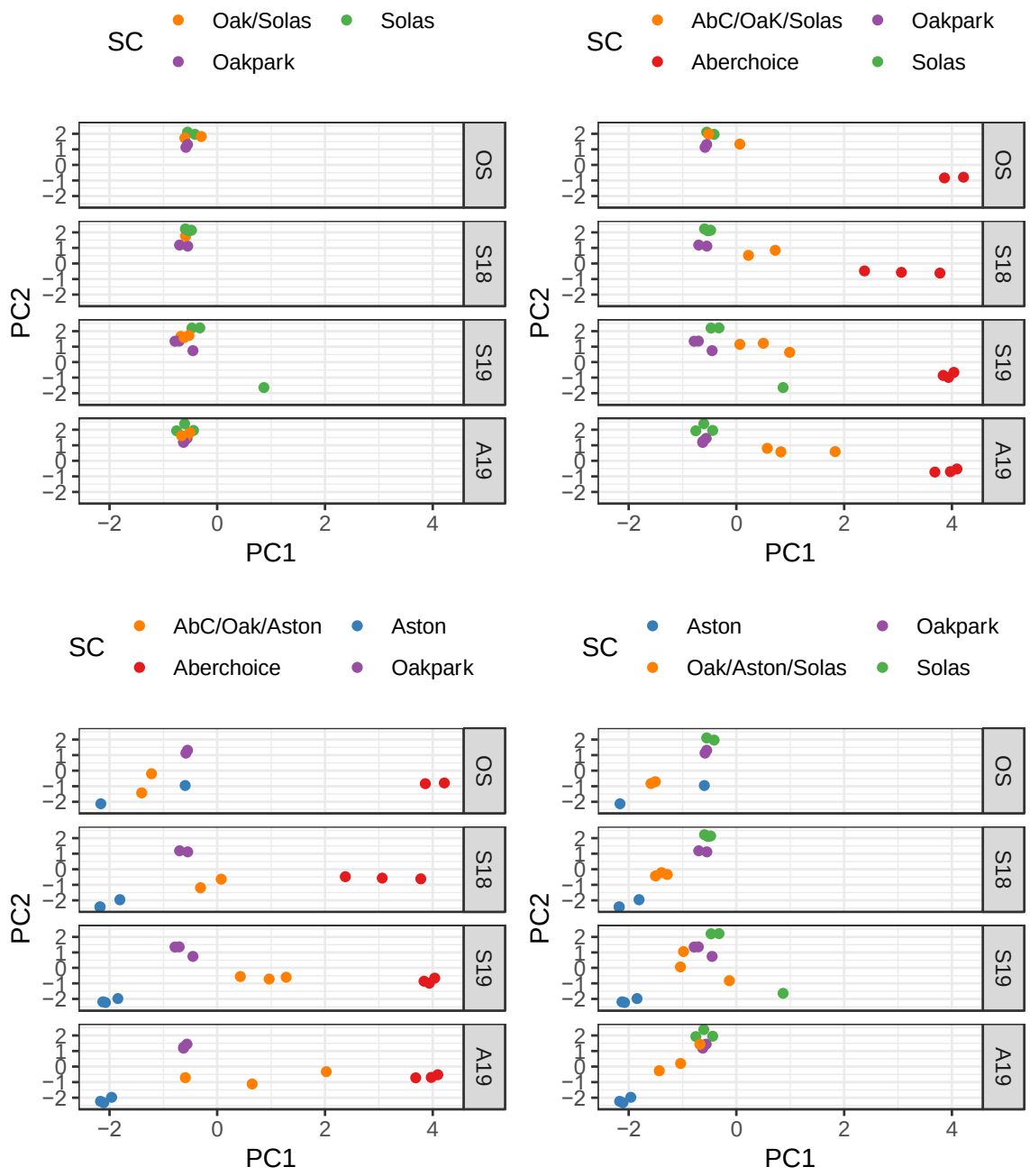
Appendix 9. The alignment rate of rate of properly paired reads to the PRG genome across 768 samples, ranked by percentage alignment.



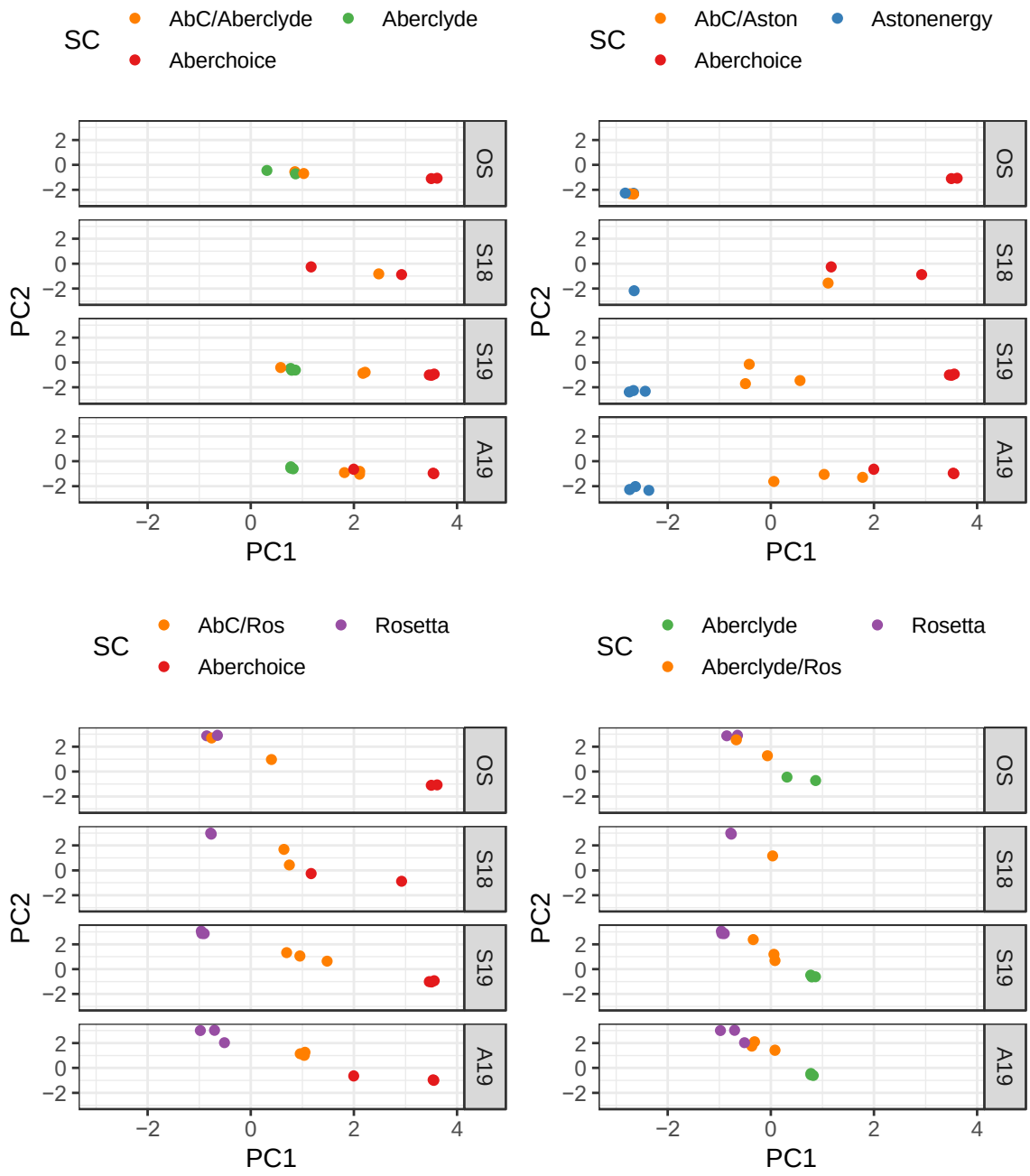
Appendix 10 The mean number of reads generated per target, in 192 targets ranked by percentage alignment.



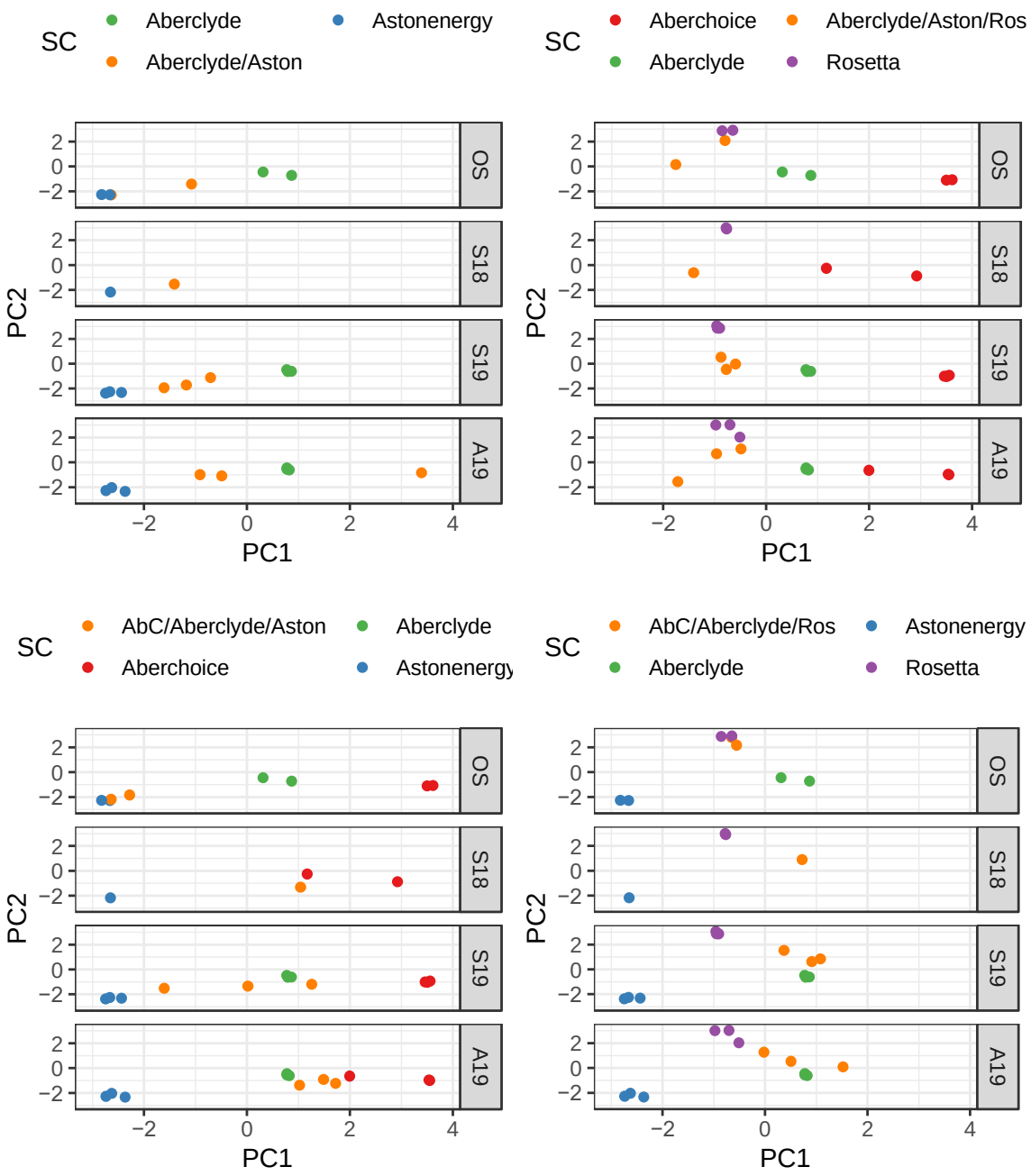
Appendix 11. PCA displaying overall trends in population structure of Study 1 two-way mixture seed combinations (SC) and the pure mixtures from which they were derived across 4 time points (OS= original seed; S18= pre-grazing spring 2018; S19 =pre-grazing spring 2019; A19 = post-grazing autumn 2019). Mixtures consist of cultivars of varying ploidy: Aberchoice (diploid, Oakpark (diploid), Astonenergy (tetraploid), Solas (tetraploid).



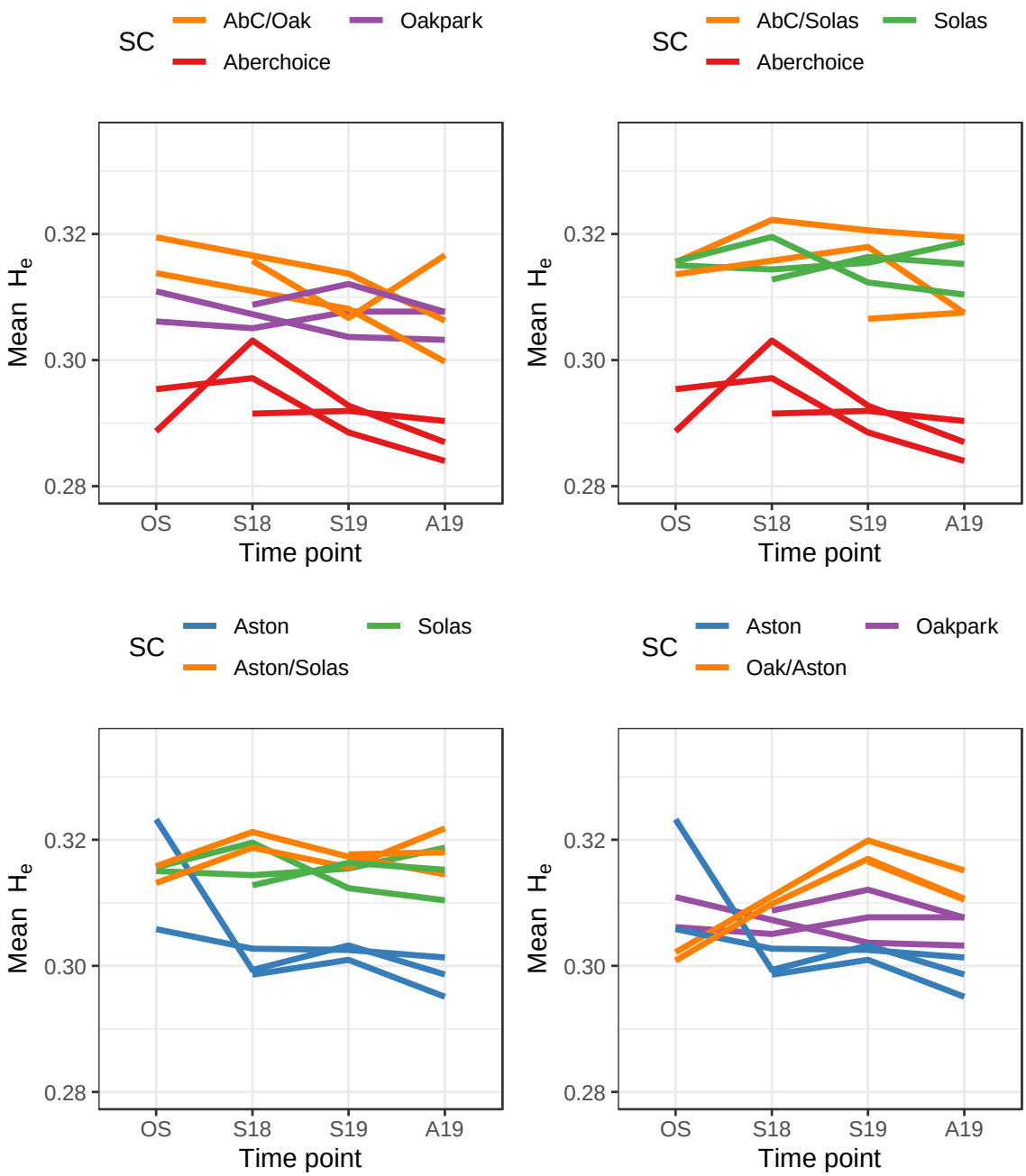
Appendix 12 PCA displaying overall trends in population structure of Study 1 two-way and three way mixture seed combinations (SC) and the pure mixtures from which they were derived across 4 time points (OS= original seed; S18= pre-grazing spring 2018; S19 =pre-grazing spring 2018; A19 = post-grazing autumn 2019). Mixtures consist of cultivars of varying ploidy: Aberchoice (diploid, Oakpark (diploid), Astonenergy (tetraploid), Solas (tetraploid).



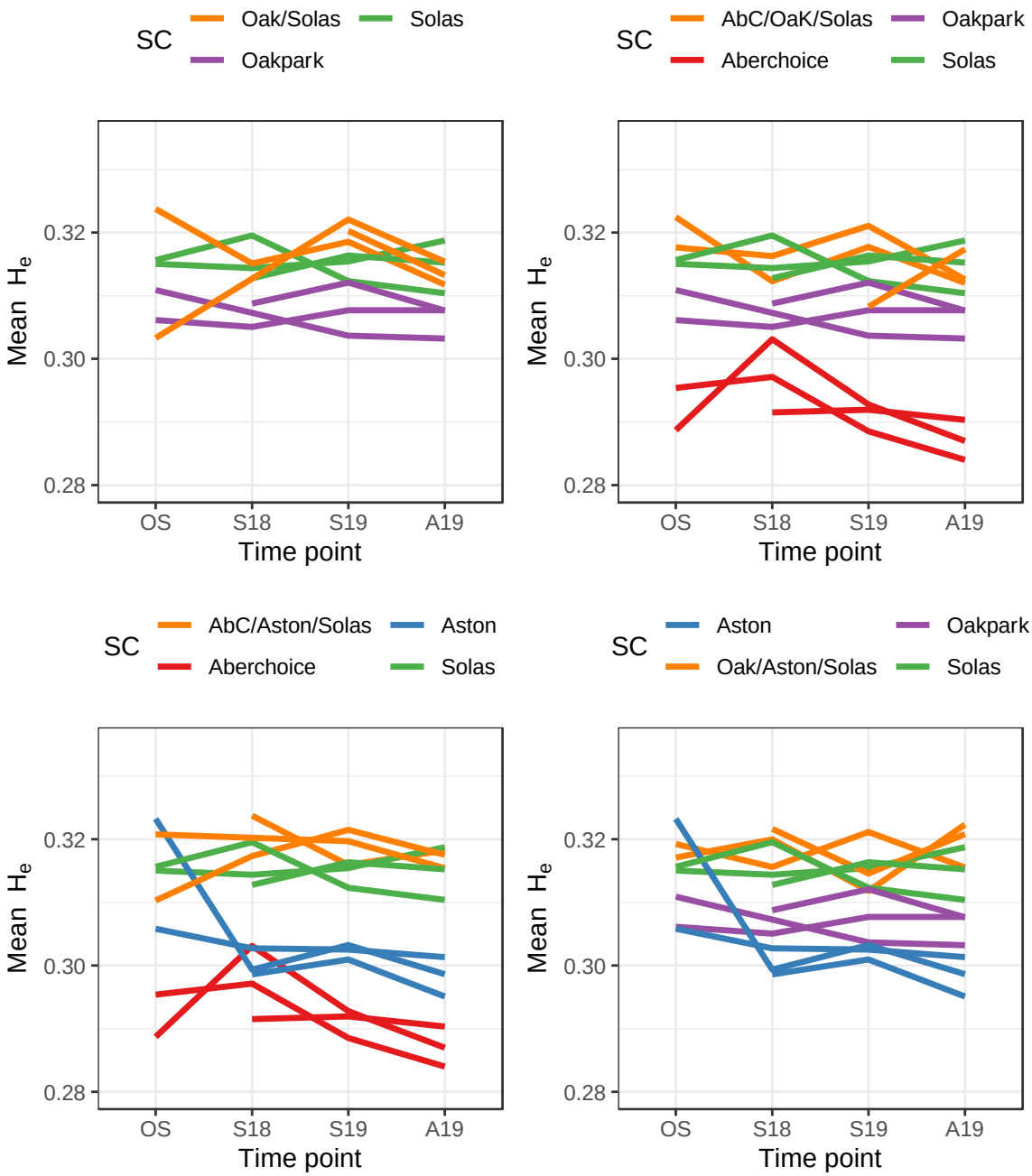
Appendix 13 Displaying overall trends in population structure of Study 2 two-way and three way mixture seed combinations (SC) and the pure mixtures from which they were derived across 4 time points (OS= original seed; S18= pre-grazing spring 2018; S19 =pre-grazing spring 2018; A19 = post-grazing autumn 2019). Mixtures consist of cultivars of varying ploidy and maturity: Aberchoice (diploid, late maturity) Aberclyde (tetraploid, intermediate maturity), Astonenergy (tetraploid, late maturity), Rosetta, (diploid, intermediate).



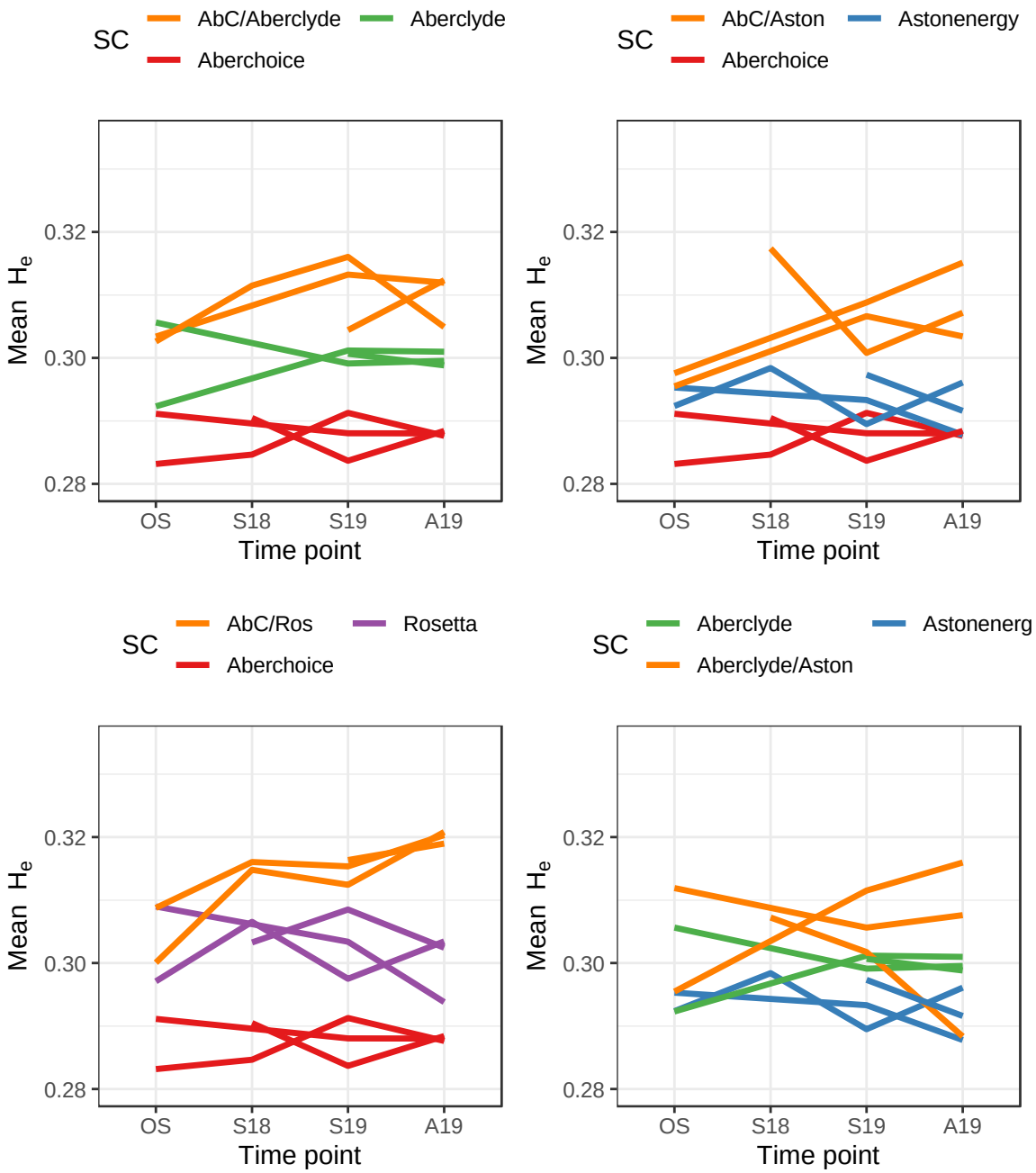
Appendix 14 PCA displaying overall trends in population structure of Study 2 two-way and three way mixture seed combinations (SC) and the pure mixtures from which they were derived across 4 time points (OS= original seed; S18= pre-grazing spring 2018; S19 =pre-grazing spring 2018; A19 = post-grazing autumn 2019). Mixtures consist of cultivars of varying ploidy and maturity: Aberchoice (diploid, late maturity) Aberclyde (tetraploid, intermediate maturity), Astonenergy (tetraploid, late maturity), Rosetta, (diploid, intermediate).



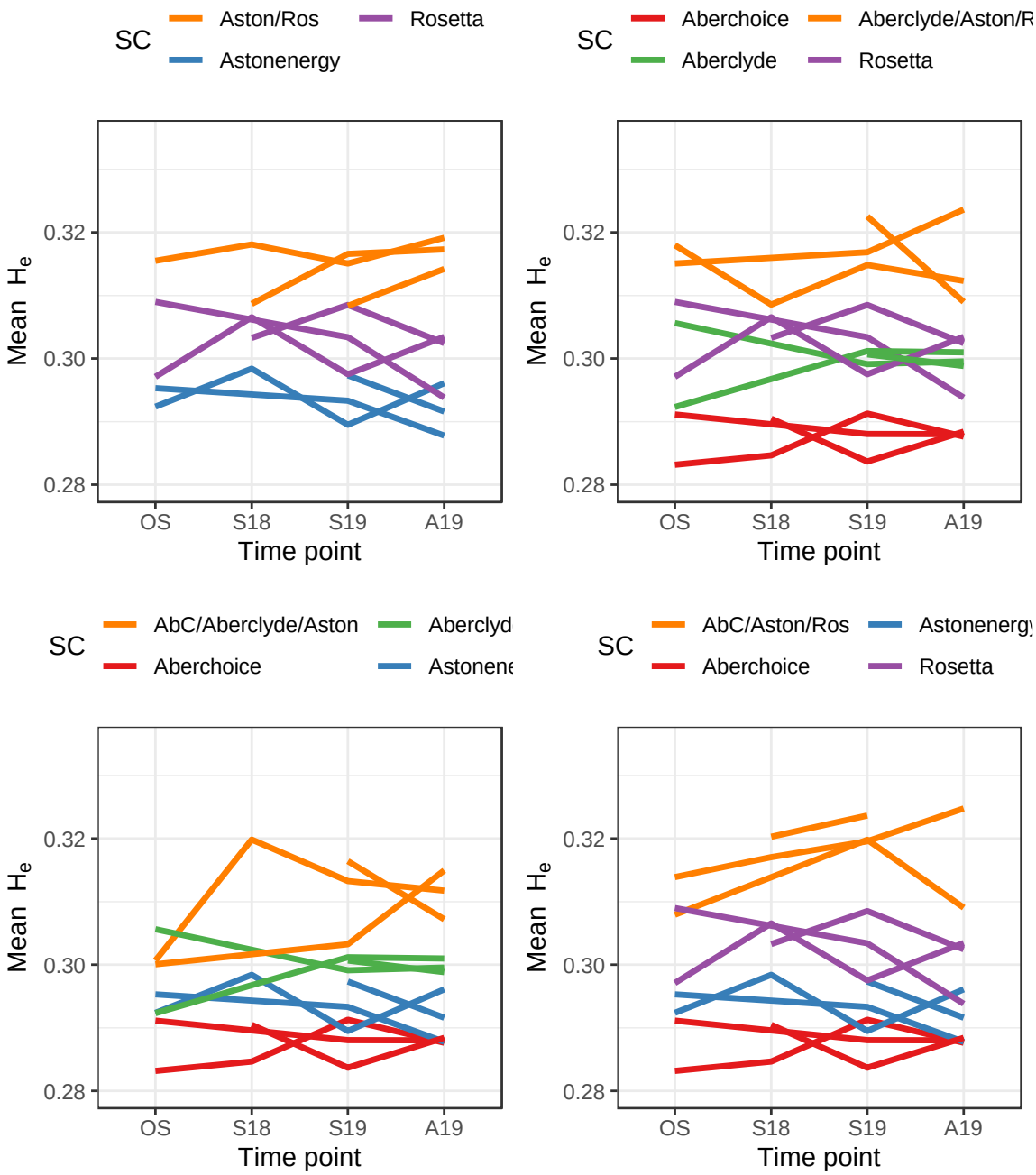
Appendix 15 Trends in mean H_e over time in Study 1 two-way mixtures across 4 time points (OS= original seed; S18= pre-grazing spring 2018; S19 =pre-grazing spring 2018; A19 = post-grazing autumn 2019). Mixtures consist of cultivars of varying ploidy: Aberchoice (diploid, Oakpark (diploid), Astonenergy (tetraploid), Solas (tetraploid).



Appendix 16. Trends in mean H_e over time in Study 1 two-way mixtures across 4 time points (OS= original seed; S18= pre-grazing spring 2018; S19 =pre-grazing spring 2018; A19 = post-grazing autumn 2019). Mixtures consist of cultivars of varying ploidy: Aberchoice (diploid, Oakpark (diploid), Astonenergy (tetraploid), Solas (tetraploid).



Appendix 17 Trends in mean H_e over time in Study 2 two-way mixtures across 4 time points (OS= original seed; S18= pregrazing spring 2018; S19 =pregrazing spring 2018; A19 = postgrazing autumn 2019). Mixtures consist of cultivars of varying ploidy and maturity: Aberchoice (diploid, late maturity) Aberclyde (tetraploid, intermediate maturity), Astonenergy (tetraploid, late maturity), Rosetta, (diploid, intermediate).



Appendix 18 Trends in mean H_e over time in Study 2 two-way mixtures across 4 time points (OS= original seed; S18= pregrazing spring 2018; S19 =pregrazing spring 2018; A19 = postgrazing autumn 2019). Mixtures consist of cultivars of varying ploidy and maturity: Aberchoice (diploid, late maturity) Aberclyde (tetraploid, intermediate maturity), Astonenergy (tetraploid, late maturity), Rosetta, (diploid, intermediate).